

STIC-Biotech/ChemLib

73150

From: Chan, Christina
Sent: Tuesday, August 13, 2002 10:44 AM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 09/697206

Please rush. Thanks Chris

-----Original Message-----

From: Davis, Minh-Tam
Sent: Tuesday, August 13, 2002 10:37 AM
To: Chan, Christina
Subject: Rush search request for 09/697206

Please search in commercial database and in issued patent files:
SEQ ID NO:1.

Thank you.

MINH TAM DAVIS

ART UNIT 1642, ROOM 8A01, MB 8E12
305-2008

porter 10/94

Point of Contact:
Mona Smith
Technical Information Specialist
CM1 6A01
Tel: 308-3278

(1)

STIC/Biotech/ChemLib
(STIC)

AUG 13 2002

RECEIVED

Searcher: M. Smith
Phone: _____
Location: _____
Date Picked Up: 8/13/02
Date Completed: 8/14/02
Searcher Prep/Review: 5
Clerical: _____
Online time: 7

TYPE OF SEARCH:

NA Sequences: 1
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Point of Contact
Mona Smith
Technical Information Specialist
OMI 8401
Tel. 202-512-2000

Query Match 65.8%; Score 2360; DB 6; Length 2731;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 2659; Conservative 0; Mismatches 10; Indels 273; Gaps 2;

```
QY 656 cagcagggttacatggcagcttcctcaagatcaagatattgcaagatttctcaaggagactca 715
Db 16 CGCAGGTTTACCATGCGAGCTTTCAGATCAAGATATTGCAAGATTCTTCAAGGACTCA 75
QY 716 atattgcaaggaggtgcagcactttgtctgaatgcctcagggtcgaagaaacggaag 775
Db 76 ATATTGCCAAGGAGGTGCGACACTTTGTCTGAATGCTCAGGGTCGAAGGAACGGGAAG 135
QY 776 ctctgggttagtttgaagtgaaggagcaccgagacctagcactacagaggcacaacaatc 835
Db 136 CTCTGGTTAGTTGTAAAGTGAGGAGCACCAGACCTTAGCACATACAGAGGCACAAACATC 195
QY 836 acatggggaccgggtattgaggtttacaaagcaacaggtgaagatttcccttaaaattg 895
Db 196 ACATGGGACCGGTATTTCAGGTTTACAAAGCAACAGGTGAAGATTTCCTTAAATTTG 255
QY 896 ctggtgtacttccaatgaggtagccagtttctctccaaggaaatcaagtcattgttc 955
Db 256 CTGGTGTACTTCCAAATGAGGTAGCCAGTTTCTCTCCAAGGAAATCAAGTCATTGTCC 315
QY 956 qcatggggggctccctttcacggccacagctgaagaagtgggtggccttcttggacagc 1015
Db 316 GCATGCGGGGCTCCCTTTTCAGGGCCACAGCTGAGAAAGTGGTGGCCTTCTTTGGACAGC 375
QY 1016 attgcccctattactgggggaaaggagccatcctcttgcacctaccagatggtaggc 1075
Db 376 ATTGCCCTATTACTGGGGGAAAGGAGGCATCCTCTTTGTCACTACCCAGATGGTAGGC 435
QY 1076 caacagggagcgtttgtcctctttgtcgtgaggaatatgcagaaatgcgttgagga 1135
Db 436 CACAGGGGAGCGTTTGTCTTCCCTTGTGAGGAATATGCACAGAATGCGTTGAGGA 495
QY 1136 agcataaagacttgggttaaagaacacattgaactcttcaggagcacagcagctgaag 1195
Db 496 AGCATAAAGACTTGTGGGTAAAGATACATTGAACCTTTCAGGAGCAGACAGCTGAAG 555
QY 1196 ttcagcagggtgcgaatcgatbtctcggccctctcatctccactccaccctccca 1255
Db 556 TTCAGCAGGTGCTGAATPCGATTCTCCTCGGCCCTCTCATTTCCACTTCCAACCCCTCCCA 615
QY 1256 ttattccagttactaccacgaatttggccctcacaatgtagagactgtatacgc 1315
Db 616 TTATTCAGTACTACCTFCAGCAATTTGTGCCCTCTACAAATGTTAGAGACTGTATAGGCC 675
QY 1316 ttcgaggtcttccctatgcagccacaattgaggacatcctggatttctcggggagttcg 1375
Db 676 TTCGAGGCTCTCCCTATGCGAGCCACAAATTCAGGACATCTCTGGATTCTCTGGGGAGTTCC 735
QY 1376 ccacagatatctgtaetcatggygtttcaatggttttgaatcacccaggggcccccacag 1435
Db 736 CCACAGATATTCTGACTCATGSGGTTTCACATGGTTTTGAATCACCAGGGCGGCCCATCAG 795
QY 1436 gqatgacctttatccagatgaagtctcgagacagacattttagctgcacagaagtgc 1495
Db 796 GAGATGCTTTATCCAGATGAAGTCTGCGGACAGAGCATTTATGGCTGCGACAGAAGTGC 855
QY 1496 ataaaaaaacatgaaggacagatatgttgaagctttcagttcagttcagctgagagatga 1555
Db 856 ATAAAAAAACATGAAGGACAGATATGTTGAAGCTTTTCAAGCTTTTCAGCTGAGGAGATGA 915
QY 1556 actttgtttaattggggggcactttaaatcgaaatggcttatcccccacag - - - - - 1605
Db 916 ACTTTGTGTTAAATGGGGGACACTTAAATCGAAATGGCTTATCCCCACCGCCCATGTAAGT 975
QY 1606 --cattgcctctctccctctcacatttccagctcctcgtcagttattcctacag 1663
Db 976 TACCATGCCCTGCTCTCTCCCTCTACACATTTCCAGCTCTCTGTCGAGTTATTTCCTACAG 1035
```

```
QY 1664 aagctgccatttaccagccctctgtgattttgaatccacagcagcactgcagccctccacag 1723
Db 1036 AAGCTGCCATTACCAGCCCTCTGTGATTTTGAATCACCAGCACATGTCAGCCCTCCACAG 1095
QY 1724 cgtactaccagcaggcactcagctcttcatgaactacacagcgtactataccacagccccc 1783
Db 1096 CGTACTACCCAGCAGGCACATCAGCTCTTCATGAATTAACAGCGTACTATCCC - - - - - 1148
QY 1784 caggttcgcctaatagtcttggctacttccctacagctgctaattcttagcgggtgcctc 1843
Db 1149 - - - - - 1148
QY 1844 cacagcctggcacggtggttcagaatgcaggcctggcctacaatactagggttaagaaa 1903
Db 1149 - - - - - 1148
QY 1904 ttcttaactcttccaaaggttaccagttaccagtatgcaaccggaggtggaacttatcacacaaatg 1963
Db 1149 - - - - - 1148
QY 1964 accagccaggactctaccacaaagaatgggttgattttaaggggcccccagcagttagaac 2023
Db 1149 - - - - - 1148
QY 2024 atcctcagaaaaagaagtgtttgaaagatgtatggtgatcttgaacacctccagacacaaga 2083
Db 1149 - - - - - ACTGTTTGAAGATGTATGGTGATCTTGAACCTCCAGACACAAGA 1194
QY 2084 aaactctagcaaatccaggggaagttgtctacactcagcgtgcagtatattccagcaaa 2143
Db 1195 AAACTTCTAGCAAAATTCAGGGGAAGTTTGTCTACACTCAGGCTGCAGTATTTTTCAGAAA 1254
QY 2144 ctgattggacaaacgggcctgtgccttattcttttggtggagtgaaaaattggagctag 2203
Db 1255 CTGATTGGACAAACGGGCGCTGTGCCCTTATCTTTGGTGAGTGAAAAAATTTTGAGCCAG 1314
QY 2204 tgaagccaaatcgttaactacagcaagcagcagcatgcagcatcacctggctcttctgtgattg 2263
Db 1315 TGAAGCCAAATCGTAACCTTACGCAAGCAGCAGCATGCAGCATACCTGGCTCTTTGCTGATTG 1374
QY 2264 caaataggcaattcaaatgtgaatttggaaatcagatgctccattacttccagttaaagt 2323
Db 1375 CAAATAGGCATTTAAAATGTGAATTTTGGAAATCAGATGCTCTCCATTACTTCCAGTTAAAGT 1434
QY 2324 ggcataagtggtttccctaaagttttaaagcttggataaaaaactccaccagtgcttacc 2383
Db 1435 GGCATCATAGGCGTTTCTCTAAGTTTTTAAGTCTTGGATAAAAAATCCACCAGTGTCTACCA 1494
QY 2384 tctccaccatgaaactcgtttaaggaaagcttcattttttgtatatattcccgctctttctctt 2443
Db 1495 TCTCCACCATGAACCTCTGTTAAGGAAGCTTCAATTTTGGTATATTCCGCTCTTTTCTCTT 1554
QY 2444 cattccctgtctctgcataatcatgcctctcttgcataagtaattcaagcataagatcctt 2503
Db 1555 CATTTCCCTGCTCTCTGCATAAATCATGCECTTCTTGTCTTAAAGTAATTCAGGATAAAGTCTT 1614
QY 2504 ggaataaaaaatcacaatcttaggagaaagaataaaattgttttttcccaagtcctctg 2563
Db 1615 GGAATAATAAATACAAATCTTAGGAGAAAGATAAATTTGTTATTTTCCAGTCTCTTG 1674
QY 2564 gccatgatgatattcttatgattaaaaacaaattaaaaattttaaaccacctgaagataaatt 2623
Db 1675 GCCATGATGATATCTTATGATTAAAAACAATAATTAATTTTAAACACCTCAAGATATATT 1734
QY 2624 agaagaaattgtgcacctccacaaacatacaaaagtttaaaagtttggatctttttctc 2683
Db 1735 AGAAGAAATTTGTCCACCCCTCCACAAACATACAAAGTTTAAAAAGTTTGGATCTTTTCTC 1794
QY 2684 agcaggtatcagttgtaataatgaattagggccaaaaatgcaaaacgaaaaatgaagca 2743
Db 1795 AGCAGGTATCAGTTGTAATAATGAATTAGGGGCCAAAAATGCAAAACGAAAAATGAAGCA 1854
```

Qy 2744 gctacatgtagttagtaattcttagttgaactgaatgaattgaatttggtgcttcatatgt 2803
 Db 1855 GCTACATGAGTGTAGTAAATTTCTAGTTGAACGTAAATGAATATGCTTGGCTTCATATGT 1914
 Qy 2804 attattttatattgtaatttttcaattatgatggtttgacttttaataagagaaattcc 2863
 Db 1915 ATTATTTTATATTGATGCTTTTTCATTTATTGATGCTTTGAGCTTTAATAAGAGAAATCC 1974
 Qy 2864 atagtttttaataatccagagtgagacaatttgaaacagtgtattctagaaaaacaataca 2923
 Db 1975 ATAGTTTTTAATATCCAGAGTGAGACAATTTGAACAGTGTATCTAGAGAAACAATACA 2034
 Qy 2924 ctaactgaacagagtggaatgcttatatatattatgatagaccttaaacctttttccctcta 2983
 Db 2035 CTAAGTGAACAGAGTGAATGCTTATATATATATATATATAGCTTTAAACCTTTTCCCTCTA 2094
 Qy 2984 atgctttaactgtcaataattataaacttttaaacgcataggaactatagtcagcatgcta 3043
 Db 2095 ATGCTTTAACTGTCAATAATATATACTTTAAAGCATAGGACTATAGTCAGCATGCTA 2154
 Qy 3044 gactgagaggttaaacactgatgcaattagaaacaggtactgatgctgcagtgtttaaacac 3103
 Db 2155 GACTGAGAGGTAAACACTGATGCAATTTAGAACAGTACTGATGCTGTCAGTGTTAACAC 2214
 Qy 3104 tatgtttagctgtgtttatgctataaaagtgcgaattatagacactagctagctagctgc 3163
 Db 2215 TATGTTTATGCTGTGTTTATGCTATATAAAGTGCAATATTAGACACTAGCTAGTACTGCTGC 2274
 Qy 3164 ctcattgaactccaaagaaacaggattttcattaaagtgcattgaattggaattggaattttctct 3223
 Db 2275 CTCATGTAACCTCCAAAGAAACAGGATTTTATTAAAGTGCATTTGAATGTGGCTATTCTCT 2334
 Qy 3224 aagttaactcatttctgtcttgatgaatgaatgcgcgtgcagattttatagagctgctat 3283
 Db 2335 AAGTACTCATATTTGCTTGTGTAATGCAATGCCGTGCAGATTTATGTGGCTGCTAT 2394
 Qy 3284 tttattttctgtgcatttaacttaacaccttaaaaggagagcaaacattttcccttctca 3343
 Db 2395 TTTTATTTTCTGTGCATTTACTTTAAACACTTTAAAGGGAGAGCAACATTTCCCTTCTCA 2454
 Qy 3344 gctgactggcaatggcccttttaactgaataggaagaaaaaagggtttgtgtgaaa 3403
 Db 2455 GCTGACTGGCAATGGCCCTTTAACTGCAATAGGAAGAAAAAAGGTTGTGTGAAA 2514
 Qy 3404 attggtgataactggcacttaagatcgaaagaaatttctgtatactgtatgctgcttaaga 3463
 Db 2515 ATTGGTGATACACTGGCACTTAAGATCGAAGAAATTTCTGTATACCTGATGCTTTAAGA 2574
 Qy 3464 tgccaaagctgccaaagctctgaaagactttaagatagggcagtaagtcttactacaat 3523
 Db 2575 TGCCAAAGCTGCCAAAGCTCTGAAAGACTTTAAAGATAGGCAGTAATGCTTACTACAAT 2634
 Qy 3524 actactgagttttgtagagtttaacatttgataataaaacttgccctgttaactcctcaaa 3583
 Db 2635 ACTACTGAGTTTTGTAGAGTTTACATTTGATATAAATAAACTTGCCCTTTAATCTCCAAA 2694
 Qy 3584 aa 3585
 Db 2695 AA 2696
 RESULT 2
 LOCUS AK000178 2140 bp mRNA linear PRI 22-FEB-2000
 DEFINITION Homo sapiens cDNA FLJ20171 fis, clone COL09761.
 ACCESSION AK000178
 VERSION AK000178.1 GI:7020093
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens colon cDNA to mRNA, clone_lib:COL clone:COL09761.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)

AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished (2000)
 REFERENCE 2 (bases 1 to 2140)
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:conaleims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
 FEATURES
 Location/Qualifiers
 1..2140
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL09761"
 /tissue_type="colon"
 /clone_lib="COL"
 /note="cloning vector pME18SFL3"
 58..1134
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAA90992.1"
 /db_xref="GI:7020094"
 /translation="MPEYLNFEKSSSVRYCASQVEDMGNIILIMISEYPNHRSDPE RVNKFESGTCRSMELIDDDTVVARGLPWQSSDDIARFFKGLNIAKGAALCLNAQ GRNGEALRVFSEEHRLDALQRKHHMGTRYIEVYKATGEDEFLKTAGGTSNEVAQFL SKENQIVRMRGLPFTATAEEVVAFFGQHGILFTVYDGRPTGDAFVLFA CEYQNALRKHKDLCKRIELFRSTAAEVQVQVLRFSAPLPIPLPTPIIPVLPQ FVPTNVYRDCIRLGLPYAATIEDILDFLGEFATDINTGHVHMLNHQGRPSGDAFLQ MKSADRAFMAAQKCHKKKHEGQIC"
 BASE COUNT 606 a 481 c 464 g 589 t
 ORIGIN
 Query Match 54.8%; Score 1962.8; DB 9; Length 2140;
 Best Local Similarity 94.2%; Pred No. 0;
 Matches 2120; Conservative 0; Mismatches 7; Indels 123; Gaps 3;

Qy 394 gaattcaagaagaattgctgctggttcacctgatattgacaaactggacgttgcacaaatg 453
 Db 1 GAATTCAGAAGAAATGTTGCCCTGGTTCACCTGATATTGACAAACTGGACGTTGCCACAATG 60
 Qy 454 acagagatttttaattttgagaagagtagttcagtcctctcgatatggagcctcacaagt 513
 Db 61 ACAGAGATTATTAATTTTGAAGAAGAGTAGTTTCAGTCTCTCGATATGGAGCCTCTCAAGTT 120
 Qy 514 gaagatatggggaataataatttttagcaatgatttcagagccttataatcacaggttttca 573
 Db 121 GAAGATATGGGAATATAATTTTAGCAATGATTTTCAGAGCCTTATAATCAGAGTTTCA 180
 Qy 574 gatccagagagtggaattacaagtttgaaagtggaaacttcagacaagatggaacttatt 633
 Db 181 GATCCAGAGAGTGAATACAAAGTTTGAAGTGGAACTTCGACAGCAAGATGAACATTATT 240
 Qy 634 gatgataaacccgtagtcaggggcagaggtttaccatggcaggttcacagtcagatcagaatatt 693
 Db 241 GATGATAACACCGTAGTCAGGGCAGGAGTTTACCATGGCAGTCTTCAGATCAAGATATT 300
 Qy 694 qcaagattcttcaagagactcaattatgccaaaggaggtgcagcactttgtcctaagct 753
 Db 301 GCAAGATTCTTCAAGGACTCAATATTGCCAAGGAGGTGCAGCACTTTGTCTGAATGCT 360
 Qy 754 cagggtcgaaggaaacggagaagcgtggttagttgttaagtgaggagcaccgagaccta 813

Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,
 Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeArcellano, K., Depayre, E., Devon, K., Dewar, K.,
 Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C.,
 Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
 Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
 Karatas, A., Lehotsky, J., Lieu, C., Locke, K., Macdonald, P.,
 Marquis, N., McEwan, P., McGuck, A., McKernan, K., McLaughlin, J.,
 Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
 Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
 Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
 Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
 Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
 Direct Submission
 Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 10, 2000 this sequence version replaced gi:7547130.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L2303
 Center clone name: lil_D13
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 132474 bases at least Q40
 Consensus quality: 144968 bases at least Q30
 Consensus quality: 150672 bases at least Q20
 Insert size: 182000; agarose-fp
 Insert size: 156233; sum-of-contigs
 Quality coverage: 3.0 in Q20 bases; agarose-fp
 Quality coverage: 3.5 in Q20 bases; sum-of-contigs

TITLE
 JOURNAL
 COMMENT

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 38 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 532: contig of 532 bp in length
 * 533 632: gap of 100 bp
 * 633 1684: contig of 1052 bp in length
 * 1685 1784: gap of 100 bp
 * 1785 2846: contig of 1062 bp in length
 * 2847 2946: gap of 100 bp
 * 2947 4075: contig of 1129 bp in length
 * 4076 4175: gap of 100 bp
 * 4176 5462: contig of 1287 bp in length
 * 5463 5562: gap of 100 bp
 * 5563 6755: contig of 1193 bp in length
 * 6756 6855: gap of 100 bp
 * 6856 8009: contig of 1154 bp in length
 * 8010 8109: gap of 100 bp
 * 8110 9599: contig of 1490 bp in length
 * 9600 9699: gap of 100 bp
 * 9700 10876: contig of 1177 bp in length
 * 10877 10976: gap of 100 bp
 * 10977 12499: contig of 1523 bp in length
 * 12500 12599: gap of 100 bp
 * 12600 13603: contig of 1004 bp in length
 * 13604 13703: gap of 100 bp
 * 13704 15007: contig of 1304 bp in length
 * 15008 15107: gap of 100 bp

* 15108 16547: contig of 1440 bp in length
 * 16548 16647: gap of 100 bp
 * 16648 17884: contig of 1237 bp in length
 * 17885 17984: gap of 100 bp
 * 17985 19796: contig of 1812 bp in length
 * 19797 19896: gap of 100 bp
 * 19897 21294: contig of 1398 bp in length
 * 21295 21394: gap of 100 bp
 * 21395 22935: contig of 1541 bp in length
 * 22936 23035: gap of 100 bp
 * 23036 26495: contig of 3460 bp in length
 * 26496 26595: gap of 100 bp
 * 26596 27849: contig of 1254 bp in length
 * 27850 27949: gap of 100 bp
 * 27950 45165: contig of 17216 bp in length
 * 45166 45265: gap of 100 bp
 * 45266 47264: contig of 1999 bp in length
 * 47265 47364: gap of 100 bp
 * 47365 51079: contig of 3715 bp in length
 * 51080 51179: gap of 100 bp
 * 51180 54251: contig of 3072 bp in length
 * 54252 54351: gap of 100 bp
 * 54352 58073: contig of 3722 bp in length
 * 58074 58173: gap of 100 bp
 * 58174 61307: contig of 3134 bp in length
 * 61308 61407: gap of 100 bp
 * 61408 65572: contig of 4165 bp in length
 * 65573 65672: gap of 100 bp
 * 65673 69628: contig of 3956 bp in length
 * 69629 69728: gap of 100 bp
 * 69729 73243: contig of 3515 bp in length
 * 73244 73343: gap of 100 bp
 * 73344 77727: contig of 4384 bp in length
 * 77728 77827: gap of 100 bp
 * 77828 83613: contig of 5786 bp in length
 * 83614 83713: gap of 100 bp
 * 83714 89358: contig of 5645 bp in length
 * 89359 89458: gap of 100 bp
 * 89459 95382: contig of 5924 bp in length
 * 95383 95482: gap of 100 bp
 * 95483 102960: contig of 7478 bp in length
 * 102961 103060: gap of 100 bp
 * 103061 112213: contig of 9153 bp in length
 * 112214 112313: gap of 100 bp
 * 112314 125685: contig of 13372 bp in length
 * 125686 125785: gap of 100 bp
 * 125786 142030: contig of 16245 bp in length
 * 142031 142130: gap of 100 bp
 * 142131 157955: contig of 15825 bp in length
 * 157956 158055: gap of 100 bp
 * 158056 159933: contig of 1878 bp in length.
 * Location/Qualifiers
 1. .159933
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-111D13"
 /clone_lib="RPC1-11 Human Male BAC"
 1. .532
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 633. .1684
 /note="assembly_fragment"
 1785. .2846
 /note="assembly_fragment"
 2947. .4075
 /note="assembly_fragment"
 4176. .5462
 /note="assembly_fragment"
 5563. .6755
 /note="assembly_fragment"

FEATURES

source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature


```

RESULT 4
AC084253
LOCUS AC084253 91840 bp DNA linear HTG 09-MAY-2001
DEFINITION Homo sapiens clone RP11-498C11, WORKING DRAFT SEQUENCE, 5 unordered
pieces.
ACCESSION AC084253
VERSION AC084253.3 GI:13123246
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 91840)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 91840)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bida,F., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Melgrim,J., Meneus,L., Mihoa,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sugnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (18-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 25, 2001 this sequence version replaced gi:11597050.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11484
Center clone name: 498_C_11
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 90064 bases at least Q40
Consensus quality: 90922 bases at least Q30
Consensus quality: 91180 bases at least Q20
Insert size: 85000; agarose-fp
Insert size: 91440; sum-of-contigs
Quality coverage: 11.1 in Q20 bases; agarose-fp
Quality coverage: 10.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

```

* be preserved.
* 1 29791: contig of 29791 bp in length
* 29792 29891: gap of 100 bp
* 29892 31022: contig of 1131 bp in length
* 31023 31122: gap of 100 bp
* 31123 57388: contig of 26266 bp in length
* 57389 57488: gap of 100 bp
* 57489 91339: contig of 33851 bp in length
* 91340 91439: gap of 100 bp
* 91440 91840: contig of 401 bp in length.
FEATURES
Location/Qualifiers
source
1. .91840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-498C11"
/clone_lib="RPC1-11 Human Male BAC"
1. .29791
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
29892_31022
/note="assembly_fragment"
31123_57388
/note="assembly_fragment"
57489_91339
/note="assembly_fragment"
91440_91840
/note="assembly_fragment"
clone_end:rf7
vector_side:right"
BASE COUNT 26234 a 17891 c 18577 g 28684 t 454 others
ORIGIN
Query Match 40.4%; Score 1449.8; DB 2; Length 91840;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1534; Conservative 0; Mismatches 10; Indels 10; Gaps 7;
QY 2038 agtgttgaagatgtatgtgtatcttgaacacctccagacacacagaaacctcttagcaaa 2097
|||||
Db 22622 ACTGTTTGAAAGATGATGGTGATCTTGAAACCTCCAGACACAGAAACCTCTAGCAAA 22681
|||||
QY 2098 ttacgggggaagtgtgtctacactcaggctgcagctatttttcagcaaacattgattggacaaa 2157
|||||
Db 22682 TTCAGGGGGAAGTTGTTCTACACTCAGGCTGCAGTATTTTCAGCAACCTGATTGGACAAA 22741
|||||
QY 2158 cgggcctgtgccttatcttttggaggagtgaaaaaatttgagctagtgaaagccaaatcgt 2217
|||||
Db 22742 CGGCGCTGTGCTTATCTTTTGGTGGAGTGAAAAAATTTGAGCCAGTGAAAGCCAAATCGT 22801
|||||
QY 2218 aactacagcaagcagcatgcagcatccttgctctcttctgtgattgcaaataggcattta 2277
|||||
Db 22802 AACTTACAGCAGCAGCATGCAGCATACCTGGCTCTCTTGTCTGATTGCAATAGGCATTTA 22861
|||||
QY 2278 aaatgtgaatttgaatcagatgtctcattctccagtttaaaagtggagcatcatagtggt 2337
|||||
Db 22862 AAATGTGAATTTGCAATCAGATGCTCCATTACTTCCAGTTTAAAGTGSCATCATAGGCGT 22921
|||||
QY 2338 ttccatagttttaaagttcttgatataaaact---ccaccagtgctctaccatctccc-accatg 2394
|||||
Db 22922 TTCTTAAGTTTTTAAGTCTTGGATAAAAATNCCNACCAGTGTCTACCATCTCCNACCATG 22981
|||||
QY 2395 aactccttgaaggagcttcattttgtatatatcccgcctctttctctcttctctccctgt 2454
|||||
Db 22982 AACNCTGTTTAGGAAGCTTCATTTTTCGTATATATCCCGCTCTTTTCTCTTCATTTCCCTGT 23041
|||||
QY 2455 cttctgcataatcatgcctctcttctgtaagttaattcaagcataaagatcttggaaataataa 2514
|||||
Db 23042 CTTCTGCATAATCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 23101
|||||
QY 2515 atcacatcttaggagagaaagaataaaattgttatttcccagtgctcttggccatgatgat 2574
|||||
Db 23102 ATCACAATCTTAGGAGAAAGAAATAAATTTGTTATTTTCCCGAGTCTCTTGGCCATGATGAT 23161
|||||

```

QY	2575	atcttatgataaaacaaataataattttaaacacctgaagataaaattagaagaattg	2634
Db	23162	ATCTTATGATTAAACAAATAATTAATTTTAAACACCTGAAGATATATTAGAAGAAATTG	23221
QY	2635	tgaacctcc-acaaaacatacaaaagttaaagaatttggaatctttttctcagcaggtatc	2693
Db	23222	TGCACCTCCNACAAACACATACAAAGTTTAAAGTTTGGATCTTTTCTCAGCAGGTATC	23281
QY	2694	agttgtaataatgaattgagggcccaaaatgcaaaacgaaataagcagctacatgta	2753
Db	23282	AGTTGTAATAATGAATTAGGGCCAAATGCAAAACGAAATAATGAAGACGTACATGTA	23341
QY	2754	gttagtaattcttagttgaaactgtaattgaatatgtgtgcttcacatgtattatttat	2813
Db	23342	GTTAGTAATTTCTAGTTGAACGTAAATCAATATTGTGGCTTCATATGATTATTAT	23401
QY	2814	atlgtaacttttttcattatgattggttggaactttaataagagaataatccatgatttta	2873
Db	23402	ATTGTACTTTTTTCATTTATTTGATGGTTTGGACTTTAATAAGAGAAATTCONTAGTTTTTA	23461
QY	2874	atccccagaagtgcagacaatttgaacagtgtaattcttagaaaaacatacaactaactgaac	2933
Db	23462	ATATCCCGAAGCTGAGACAAATTTGACACAGTGTATTTCTAGAAAAACAATACACTAAGTGAAC	23521
QY	2934	agaagtgaatgccttatatatattatgatagcccttaaccccttttcccttaatgccttaac	2993
Db	23522	AGAAGTGAATGCTATATATATTATGATAGCCTTAAACCTTTTTCCTCTAATGCCTTAAC	23581
QY	2994	tgcaaaataattaaacctttttaaagcataggaactatagtcacatgcagctagagagg	3053
Db	23582	TGTCAAAATAATTAAACCTTTTAAAGCATAGGACTATAGTCAGCATGCTAGACTGAGAGG	23641
QY	3054	taaacactgatgcaattagacaggtactgatgctgcagtggtttaaactatgttttagc	3113
Db	23642	TAAACACTGATGCAATTTAGAACAGGTACTGATGCTGTCAGTGTTTAACACTATGTTTAGC	23701
QY	3114	tggtgttatgctataaaagtgcataattagacacactagctagtactgc--tgccctcatgta	3171
Db	23702	TGTGTTTTATGCTATAAAGTGAATATTAGACACTAGTACTGCTGCTGCTATTTTTA	23761
QY	3172	actccc--aagaaacagagatttcattaaagtcgatgaattggtatatttctctaagtta	3229
Db	23762	ACTNCCAAAGNAAACAGGAATTCATTTAAGTGCATTTGAATGTGGCTATTTCTCTAAGTTA	23821
QY	3230	ctcatattgt-cccttgcttgaatgcaatgccgtgcagattttatgagctgctattttta	3288
Db	23822	CTCATATTGTCCTTGTGTAATGCAATCCCGTGCAGATTTTATGTGGCTGCTATTTTTA	23881
QY	3289	ttttc-tgtgcattactttaacaccttaaaaggagagcaaaacatttccttcttcagctg	3347
Db	23882	TTTTTCNTGTGCATTACTTTTAACACCTTTAAAGGGAGAGCAAAACATTTCCCTTCCTCAGCTG	23941
QY	3348	actggaatggccctttaactgcaataggaagaaagaaaaagggtttgtgaaaaattg	3407
Db	23942	ACTGGCAATGGCCCTTTAACTGCAATAGGAAGAAAAAAGGTTTGTGAAAAATTG	24001
QY	3408	gtgataactggcacttaagatcgaagaaagaatttctgtatctgtatgccttaagatgcc	3467
Db	24002	GTGATACTGGCACTTAAGATCGAAAGAAATTTCTGTATACTTGTATGCTTAAAGATGCC	24061
QY	3468	caaaagctgccccaaagctctgaaagactttaagataggcagtaaatgtctactacaatacta	3527
Db	24062	CAAAAGCTGCCAAAGCTCTGAAAGACTTTTAAGATAGGCAGTAATGCTTACTAGAACTACTA	24121
QY	3528	ctgaagtttttgaggttaacatttaataataaaactgctgttttaactcaaa	3581
Db	24122	CTGAGTTTTTTGTAGAGTTAACATTTTGATAAAACCTTGCCCTGTTTAATCTCAA	24175
RESULT	5		
AK001805		1624 bp	mrna
LOCUS			linear PRI 22-FEB-2000

DEFINITION	Homo sapiens cDNA FLJ10943 fis, clone OVARC1001360.									
ACCESSION	AK001805									
VERSION	AK001805.1 GI:7023305									
KEYWORDS	oligo capping; fis (full insert sequence).									
SOURCE	Homo sapiens ovarian cancer cDNA to mRNA, clone_lib:OVARC1 clone:OVARC1001360.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1 (sites) Isozaki,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masubo,Y., Ninomiya,K. and Iwayanagi,T.									
TITLE	NEDO human cDNA sequencing project									
JOURNAL	Unpublished (2000)									
REFERENCE	2 (bases 1 to 1624)									
AUTHORS	Isozaki,T. and Otsuki,T.									
TITLE	Direct Submission									
JOURNAL	Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases, Takao Isozaki, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)									
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.									
FEATURES	Location/Qualifiers									
Source	1..1624 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="OVARC1001360" /clone_lib="OVARC1" /tissue_type="ovarian cancer" /note="Cloning vector: pME18SFL3"									
BASE COUNT	498 a 320 c 288 g 518 t									
ORIGIN										
Query Match	37.0%; Score 1324.8; DB 9; Length 1624;									
Best Local Similarity	99.5%; Pred. No. 6.9e-308;									
Matches 1329; Conservative	0; Mismatches 7; Indels 0; Gaps 0;									
QY	2038	agtgtttgaaagatgtatgtgtatcttgaacccctccagacacagaacaaactcttagcaaa	2097							
Db	289	AGTGTTTGAAAGATGTATGGTGATCTTGAAACCTCCAGACACAGAAACCTTAGCAAA	348							
QY	2098	ttcaggggaagttgtctacacccagctgcagatttttcagcaacttgatggacaaa	2157							
Db	349	TTCAAGGGGAAGTTTGTCTACACTCAGCTCAGTATTTTTCAGCAAACTTGATTTGGACAAA	408							
QY	2158	cgggcctgtgccttatcttttgggtgagtgagaaaaatttgagctagtgaagcccaaatcgt	2217							
Db	409	CGGGCCTGTGCCCTATCTTTTGGTGAGTGAAAAATTTTGAGCCAGTGAAGCCAAATCGT	468							
QY	2218	aaectacagcaagcagcatgcagcataccctgcgtctcttctgtgattgcacaataggcattta	2277							
Db	469	AACTTACAGCAAGCAGCATGCAGCATACCTTGGCTCTTTGTCTGATTGCAAAATAGGCATTTA	528							
QY	2278	aaatggaatttggaatcagatgtccttacttccattctccagttaaagtggcatactagtgct	2337							
Db	529	AAATGGAATTTGGAATCAGATGCTCCCACTACTTCCAGTTAAAGTGGCATCATAGCCGT	588							
QY	2338	ttccctaaagttttaagctcttgataaaaaactcccacagtgctctaccatctccacatgaac	2397							
Db	589	TTCTTAAGTTTAAAGTCTTGATATAAAACTCCACCAGTGTCTACCATCTCCACCATGAAC	648							

QY 2398 tctgttaaggaagcttcattttgtatatcccgctctctttctcttcattccctgtctt 2457
 Db TCTGTTAAGGAAGCTCAATTCGTATATCCCGCTCTTTTCTCTTCATTCCTCTCTT
 QY 2458 ctgcataatcatgcctcttctgtaagtaattcaagcataagatcttggataataaaatc 2517
 Db CTGCATAAATCATGCCCTCTTGTGAAGTAATCAAGCATAGATCTTGGTAATAATAAAATC 768
 QY 2518 acaactttaggagaagaataaaattgttatttccagctctcttggccatgatgatatc 2577
 Db ACAATCTTAGGGAAGAATAAATGTATTTTCCAGCTCTCTTGGCCCATGATGATATC 828
 QY 2578 ttatgattaaaaacaaatttaaatttttaaacacctgaagataaaattagaagaattgtgc 2637
 Db TTATGATTAAAAACAATAATTTTAAACACCTGAAGATATATTAGAAGAAATTTGTGC 888
 QY 2638 accctccaaacatacacaaagttaaagtgttgatcttttctcagcaggtatcagtt 2697
 Db ACCCTCCACAAAACATACAAAAGTTTAAAAGTTTGGATCTTTTCTCAGCAGGTATCAGTT 948
 QY 2698 gtaataatgaattagggccaaaatgcacaaacgaaaatgaagcagctcacatgtagttta 2757
 Db GTAATAATGAATTAGGGGCCAAAATGCAAAACGAAAATGAAGCAGCTACATGTAGTTA 1008
 QY 2758 gtaattcttagttgaactgttaattgaatatattgtgcttcatatgtattatttatattg 2817
 Db GTAATTCTTAGTTGAACGTGAATGAATATGTGGCTTCATATGTATTTATTTATATTG 1068
 QY 2818 tactttttcattatgattgttgactttaataagagaattccatgatttttaaat 2877
 Db TACTTTTTTCATTATGATGGTTTGGACTTTAATAAGAAAATTCATAGTTTAAATAT 1128
 QY 2878 ccagaagtgcagacaatttgaaacagtgattcttagaaaaacaatacactaactgaacagaa 2937
 Db CCAGAAGTGAGACAATTTGAACAGTGATTTAGAAAACAATACACTAAGTGAACAGAA 1188
 QY 2938 gtgaatgcttatatatattatgatagccttaaaccttttctcctaattccttaactgtc 2997
 Db GTGAATGCTTATATATATTATGATAGCTTAAACCTTTTTCCTCTAATGCTTAACTGTC 1248
 QY 2998 aaataattaaaccttttaagcatagagactatagtcagcagctagactgagagtaaa 3057
 Db AAATAATTATAACCTTTTAAAGCATAGGACTATAGTCAGCATGCTAGACTGAGAGTAA 1308
 QY 3058 cactgatgcaattagaacaggtactgatgctgctcagtggttaaacactatgttttagctgtg 3117
 Db CACTGATGCAATTAGAACAGGTACTGATGCTGCTAGCTGTTTAACTATGTTTAGCTGTG 1368
 QY 3118 ttatgctataaaagtgcataattagacactagctagtagtactgctgcctcatgtaactcca 3177
 Db TTTATGCTATAAAGTGCAATATCAGACACTAGTACTAGTCTGCTCATGTAACCTCCA 1428
 QY 3178 aagaaaaacaggatttcaattgaattgaattgaattgaattgaattgaattgaattgaatt 3237
 Db AAGAAAACAGGATTCATTAAAGTGCAATGAATGTGCTATTTCTCTAAGTTACTCATATT 1488
 QY 3238 gtcccttgctgaatgcaatgcgctgacatattatagcgtcgtattttttttttctgtg 3297
 Db GTCCCTTGTGCTGAATGCAATGCGGTGACATTTATGTGCTGCTATTTTATTTCTGTG 1548
 QY 3298 cattactttaaacaccttaagggagaagcaaacatttcttctcagctgactggcaatg 3357
 Db CATTACTTTAAACACCTTAAAGGAGAAGCAACATTTCTTCTCAGCTGACTGGCAATG 1608
 QY 3358 gccctttaactgcaat 3373
 Db GCCCTTTAACTGCAAT 1624

RESULT 6
 AK025571 3406 bp mRNA linear PRI 29-SEP-2000
 LOCUS AK025571
 DEFINITION Homo sapiens cDNA: FLJ21918 fis, clone HEP04006.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM

REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT

FEATURES
 source

CDS

BASE COUNT
 ORIGIN

Query Match 17.9%; Score 640.8; DB 9; Length 3406;
 Best Local Similarity 60.2%; Pred. No. 2.7e-143;
 Matches 1215; Conservative 0; Mismatches 722; Indels 81; Gaps 6;

QY 1 ctttttggatcactgctggggccaccggggccaaagcttagctcgatgagaaggagttg 60
 Db 184 CTCTTCGGGGCTACGGCGGGTGGCTGGGACGGGACCTGGGCTGGGACGCGACTTA 243
 QY 61 atctcgtctgttgaaagtctggtatctggtccaaagaaggtggacagtgcacgaa 120
 Db 244 ATCTCTCTAGTTTGGCAAGTGGTTGAGCGCGGAGCGCGCAGGTGGGGACGCTGCACAA 303

AK025571
 AK025571.1 GI:10438125
 oligo capping; fis (full insert sequence).
 Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
 clone:HEP04006.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
 Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
 Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 Unpublished (2000)
 2 (bases 1 to 3406)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
 Shibahara,T., Tanaka,T. and Nakamura,Y.
 Direct Submission
 Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing; Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

Location/Qualifiers
 1..3406
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP04006"
 /cell_line="hepg2"
 /cell_type="hepatoma"
 /clone_lib="HEP"
 /note="cloning vector pME18SFL3"
 88..2271
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15173.1"
 /db_xref="GI:10438126"
 /translation="MTTPPPPPPPPGDPDPAADPCWPGLSVLVFGATAGALGRD
 LGSDETDLILYQVVPERSROVGLIHKSLVRAEAAALSTOCREASGLSADSLARAE
 LDKVLQFSQLVNGDVALGGPGYMLCTDGOQLLRQVLPHEARKNLVLDPDFESYD
 LRREHMHQSTCPARDLTVATMAOGLGLETDDDFGVVEKTVLVLHLLKEFS
 SOLFSKPEVIKQYETGPCSDSTVPCYSSKADVDSDTSVWRAGLPWSSDDVARE
 FKGLNVARGGVALCLNAQRNGEALIRFVDSQDLALQRHKHMGVRYIEYKATG
 BEFVKTAGCTSLVARFLSREDQVILRLRGLPFSAGTDLVGLGPECVPTGTEGLL
 FVRHPDGRPTDAFALFACELAAALRRHGMKRYTELFRSTAAEYQOVLNRYAS
 GPLLTLPAPLLPIPPAPGTCRCVRLRGLPYTATIEDILSFLGCEAAADLRPHGVH
 MYLNOQGRPSGDAFIOMTSAERALAAQCHRKVMKRYEIVVPPSTESTEMRMVGGT
 LGRSGMSPPLKLPCLSPPTTFTQATPLTITPTAALYPSLLPALPAARVPAPTFVA
 YPGPATQLYLNATYPPSPVPTTGYLTPTTAALASAPTSLVLSQSGALVRMQGVP
 YTAGMKDLLSVFOAYQLPADDYTSMLPVGDPDPRTVLQAPKEWCL"
 731 a 1030 c 942 g 703 t

CORIXA CORPORATION (US)					
Location/Qualifiers					
Source	1. .503				
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
BASE COUNT	188 a	69 c	78 g	168 t	
ORIGIN					
Query Match 13.9%; Score 498.4; DB 6; Length 503;					
Best Local Similarity 99.8%; Pred. No. 4.5e-109;					
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	2572	gatatcttatgattaaaaacaataatttttaaacacacctgaagaataattagaagaaa			2631
Db	1	GATATCTTATGATTAAAAACAATAATTAAATTTTAAAAACACCTGAAGATATATTAGAAGAAA			60
QY	2632	ttgtgcacctccacaaacatacaaaagtcttaaaagtcttgatctttctcagcaggtg			2691
Db	61	TTGTGCACCCCTCCACAAACATACAAAGTTTAAAGTTTGGATCTTTTCTCAGCAGGTA			120
QY	2692	tcagttgtaataatgaattaggggcccaaatgcgaaacgaaaaatgaagcagctacatg			2751
Db	121	TCAGTTGTAAATTAATGAATTAGGGGCCAAAATGCAAAACGAAAAATGAAGCAGCTACATG			180
QY	2752	tagttagtaattctagtttgaaactgtaattgaattattgtggtcttcataatgtattattt			2811
Db	181	TAGTTAGTAAATTTCTAGTTTGAACGTAAATTCGAATATTGTGCTTCATATGTATTATTTT			240
QY	2812	atatgtacttttttcatttattgtatggtttggaactttaataagagaaaatcccatagttt			2871
Db	241	ATATTGTACTTTTTCATTTATTATTTGATGGTTTGGACTTTTAATAAGAGAAAATTCATATGTTT			300
QY	2872	taatacccagaagtgagacaatttgaacagtgattctcagaaaaacatacaactaga			2931
Db	301	TAATATCCAGAGTGAGACAAATTTGAACAGTGTAATTCAGAAAAACAATACACTAAGTGA			360
QY	2932	acagaagtgaatgcttatatatattatgatagccttaaacctttttccctcctaagcctta			2991
Db	361	ACAGAAGTGAATGCTTATATATATTATATATATGATAGCCTTAAACCTTTTCCCTCTAATGCCTTA			420
QY	2992	actgtcaataattataacaccttttaagcataggactatagtcagcatgctagactgaga			3051
Db	421	ACTGTCAATAATTAACCTTTTAAAGCATAGGACTATATAGTCAGCATGCTAGACTGAGA			480
QY	3052	ggtaaacactgatgcaatta			3071
Db	481	GGTAAACACTGATGCAATTA			500
RESULT	8				
LOCUS	AX062425				
DEFINITION	Sequence 52 from Patent WO0100828.				
ACCESSION	AX062425				
VERSION	AX062425.1 GI:12540300				
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
TITLE	Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W. and Mannion, J.				
JOURNAL	Compositions and methods for the therapy and diagnosis of lung cancer				
FEATURES	source				
Source	1. .503				
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
BASE COUNT	188 a	69 c	78 g	168 t	

ORIGIN	Query Match 13.9%; Score 498.4; DB 6; Length 503; Best Local Similarity 99.8%; Pred. No. 4.5e-109; Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2572	gatatcttatgattaaaaacaataataatttttaaacacacctgaagaataaattagagaaga 2631
Db 1	GATATCTTTATGATTAAAAACAATAATTAAATTTTAAACACCTGAAGATATATTAGAAGAAA 60
QY 2632	ttgtgcacctccacaaacatacaaaagtcttaaaagtcttgatcttttctcagcaggtg 2691
Db 61	TTGTGCACCCCTCCACAAAACATACAAAGTTTAAAGTTTGGATCTTTTCTCAGCAGGTA 120
QY 2692	tcagttgtataataatgaattaggggcccaaatgcgaaacgaaaaatgaagcagctacatg 2751
Db 121	TCAGTTGTAAATAATGAATTAGGGGCCAAAATGCAAAACGAAAAATGAAGCAGCTACATG 180
QY 2752	tagttagtaattcttagtttgaaactgtaattgaattattgtggcttcataatgtattttt 2811
Db 181	TAGTTAGTAAATTTCTAGTTTGAACGTAAATTCGAATATTGTGCTTCATATGTATTATTTT 240
QY 2812	atatgtacttttttcatttattgtatggtttggaactttaataagagaaaatcccatagttt 2871
Db 241	ATATTGTACTTTTTCATTTATTATGATGGTTTGGACTTTTAATAAGAGAAAATTCATATGTTT 300
QY 2872	taatacccagaagtgagacaatttgaacagtgattctcagaaaaacaatacaactaactga 2931
Db 301	TAATATCCAGAGTGAGACAAATTTGAACAGTGTAATTCAGAAAAACAATACACTAAGTGA 360
QY 2932	acagaagtgaatgcttatatatattatgatagccttaaacctttttccctcctaagcctta 2991
Db 361	ACAGAAGTGAATGCTTATATATATTATGATAGCCTTAAACCTTTTCCCTCTAATGCCTTA 420
QY 2992	actgtcaataataattataacaccttttaagcataggactatagtcagcatgctagactgaga 3051
Db 421	ACTGTCAATAATTAATTAACCTTTTAAAGCATAGGACTATAGTCAGCATGCTAGACTGAGA 480
QY 3052	ggtaaacactgatgcaatta 3071
Db 481	GGTAAACACTGATGCAATTA 500
RESULT 9	
AX340160	
LOCUS	AX340160 499 bp DNA linear PAT 10-JAN-2002
DEFINITION	Sequence 407 from Patent WO0196388.
ACCESSION	AX340160
VERSION	AX340160.1 GI:18136141
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Jiang, Y., Harlocker, S.L. and Secrist, H.
TITLE	Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL	Patent: WO 0196388-A 407 20-DEC-2001;
FEATURES	CORIXA CORPORATION (US)
source	Location/Qualifiers 1. .499 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	155 a 90 c 118 g 134 t 2 others
ORIGIN	
Query Match	13.5%; Score 485.2; DB 6; Length 499;
Best Local Similarity	99.6%; Pred. No. 6.8e-106;
Matches 496; Conservative	1; Mismatches 0; Indels 1; Gaps 1;

```
Qy 351 actattacctgaatgctcttattctcttttttgatctctcgaagaattcaagaatgttg 410
Db 1 ACTATTACCTGAATGCTTCTATTCTCTTTTGTATCTTCGAAAGAAATCAAGAAATGTTG 60
Qy 411 ccttggttcaacctgatattgacaaactggacgtttgcacaaatgacagagattttaaat 470
Db 61 CCCGTGGTCACTGATATTGCAAACTGGAGCGTTGCCACAATGACAGAGTA-TTAAATTT 119
Qy 471 tgagaagtagtctcagtcctcgcgatgagcctctcaagtgaagatatggggaatat 530
Db 120 TGAGAAGAGTAGTTCAGTCTCTCATATGGAGCCTCTCAAGTTCGAATATGGGGAATAT 179
Qy 531 aatttgcaaatgatttcagagccttataatcacaggttttcagatccagagagagtga 590
Db 180 AATTTTAGCAATGATTTTCAGAGCCTTATAATCACAGGTTTTCAGATCCAGAGAGAGTGA 239
Qy 591 ttacaagttgaaagtgaacttcagcaagatggaacttatttatataacacccgtagt 650
Db 240 TTACAAGTTTGAAGTGGAACTTCGACGAAGATGGAACCTTATGATGATAACACCGTAGT 299
Qy 651 cagggcacaggtttaccatggcagcttcagatcaagatatcgaagattcttcaagg 710
Db 300 CAGGCACGAGGTTTACCATGGCACTTCTCAGATCAAGATATTCAGAGTCTTCAAGG 359
Qy 711 actcaattgccaaaggagggtgcagcactttgtctgaatgctcaggtgcgaagaaagg 770
Db 360 ACTCAATATTGCCAAGGAGGTGCAGCAGTCTTGTCTGAATGCTCAGGCTCGAAGGAACGG 419
Qy 771 agaagctcgttggtttgtaagtgaaggagcaccgagacccagactagcactacagagca 830
Db 420 AGAAGCTCTGGTTAGGTTTGAAGTGAGGACCGCAGACCTAGCAGTACAGAGGCAAA 479
Qy 831 acatcacatggggaccgcg 848
Db 480 ACATCACATGGGGACCCD 497
RESULT 10
AX351465/c
LOCUS AX351465 434 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 212 from Patent WO0196390.
ACCESSION AX351465
VERSION AX351465.1 GI:18616812
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
JOURNAL Compositions and methods for the therapy and diagnosis of colon
PATENT: WO 0196390-A 212 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 434
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 112 a 114 c 73 g 135 t
ORIGIN
Query Match 12.1%; Score 434; DB 6; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.4e-93;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 470 ttgagagagtagtgcgtctcgcgatgagcctcgaatgagagatggggaata 529
Db 434 TTGAGAGAGTAGTTCAGTCTCTCATATGGAGCCTCTCAAGTTCGAAGATATGGGGAATA 375
Qy 530 taattttgacaatgatttcagagccttataatcacaggttttcagatccagagagagtga 589
Db 374 TAATTTTAGCAATGATTTTCAGAGCCTTATAATCACAGGTTTTCAGATCCAGAGAGAGTGA 315
```

```
Qy 590 attacaagtttgaagtggaaacttgcaagatggaaacttatgatgataaacaccgtag 649
Db 314 ATTACAGTTTGAAGTGGAACTTCGACGAAGATGGAACTTATTGATGATTAACACCGTAG 255
Qy 650 tcagggcacaggtttaccatggcagcttcagatcaagatatgcaagattcttcaaaag 709
Db 254 TCAGGGCACGAGGTTTACCATGGCAGTCTTCAGATCAAGATATTGCAAGATTTCTCAAAG 195
Qy 710 gactcaatattgcaaggaggtgcagcactttgtcgaatgctcaggtgcgaaggaacg 769
Db 194 GACTCAATATTGCCAAGGAGGTGCAGCAGTCTGCTCAATGCTCAGGCTCGAAGGAACG 135
Qy 770 gagaagctcgttggtttgtaagtgaaggacccagacccagactagcactacagagggaca 829
Db 134 GAGAAGCTCTGGTTAGGTTGTAAGTGAGGAGCACCGAGACCTAGCAGTACAGAGGCACA 75
Qy 830 aacatcacatggggaccgcggtatattgaggtttacaagaacacgggtgaagattctccta 889
Db 74 AACATCACATGGGACCGGTATATTGAGGTTTACAAGCAACAGGTGAAGATTTCCCTTA 15
Qy 890 aaattgctggtggt 903
Db 14 AAATTGCTGGTGGT 1
RESULT 11
AX193256
LOCUS AX193256 353 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 823 from Patent WO0149716.
ACCESSION AX193256
VERSION AX193256.1 GI:15211207
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
TITLE Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 823 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 353
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 122 a 46 c 54 g 129 t 2 others
ORIGIN
Query Match 9.8%; Score 351; DB 6; Length 353;
Best Local Similarity 99.4%; Pred. No. 1.3e-73;
Matches 351; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2663 aaaagtttgatcttttctcagcaggtatcagttgtaataatgaattaggggcaaaa 2722
Db 1 AAAAGTTTGGATCTTTTCTCAGCAGGTATCAGTTGTAATAATGAATAGGGGCCAAAA 60
Qy 2723 tgcaaacgaaaaatgaagcagctacatgtagtagtaattctagttggaactgtaatt 2782
Db 61 TGCAAAACGAAAAATGAAGCAGCTACATGTAGTAGTAATTTCTAGTTGAACGTAAAT 120
Qy 2783 gaattatggcttcacatgtattatttatatttatgtacttttttcattatgaggttg 2842
Db 121 GAATATTGGCTTTCATATGTATTATTATTATTTATTTTCTTTTCTATTATTATGAGGNTG 180
Qy 2843 gactttaataagaaaaatcccatagtttttaatatccagaaagtgcagacaatttgaacag 2902
Db 181 GACTTTTAATAGACAATTCATAGTTTAAATATATCCAGAAAGTCAGACAATTTGAACAG 240
Qy 2903 tgtattctagaaaaacaatacacactaacactgaacagagtgaaatgcttatattatgata 2962
```

```
Db 241 TGTATTCTAGAAAACATACACTAACTGAACAGAGTGAATGCTTATATATATATNATA 300
QY 2563 gccttaaaccttttctctaaagccttaactgtcaaatataacatttt 3015
Db 301 GCCTTAAACCTTTTCTCTAATGCCCTTAAGTGTCAAAATAATATACCTTTT 353

RESULT 12
AX335697/c 377 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 6206 from Patent WO0194629.
ACCESSION AX335697
VERSION AX335697.1 GI:18126416
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Hortigan,S., Soppet,D.R. and Weaver,Z.
Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6206 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
SOURCE Location/Qualifiers
1. .377
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 122 a 70 c 66 g 118 t 1 others
ORIGIN

Query Match 9.8%; Score 350.4; DB 6; Length 377;
Best Local Similarity 98.9%; Pred. No. 1.8e-73;
Matches 373; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 3205 tgaatgtggatatt-ctctaagttactcatattgt-cccttgcctgaatgcaatgccgt 3262
Db 377 TGAATGTGGATATTTCCTTAAGTTACTCATATTGNCCTTTGCTTGAATGCAATGCCGT 318
QY 3263 gcagatttatgagctgctatttttttctgtgcattactttaacaccttaaaaggag 3322
Db 317 GCAGATTATGAGCTGCTATTTTATTCTGTGCATTACTTTAACACCTTAAAGGGAG 258
QY 3323 aagcaaacatttctcttcagctgactgccaatggccctttaaactgcaataggaagaaa 3382
Db 257 AAGCAAAACATTTCTCTTCAGCTGACTGGCAATGGCCCTTTAACTGCAATAGGAAGAAA 198
QY 3383 aaaaaaagggttgtgtaaaattggtgataactggcacttaagatcgaaaagaaatttc 3442
Db 197 AAAAAAAGGTTTGTGGAATTTGGTGATTAACCTTGAAGTCAAGATCGAAAAGAAATTC 138
QY 3443 tgtatacttgatgctttaaagtgcccaaaagctgcccacactgctgaaagactttaagata 3502
Db 137 TGTATATTGATGCTTAAGATGCCCCAAGCTGCCAAAGCTCTGAAAGACTTTAAAGATA 78
QY 3503 ggcagtaactcttactacaactactactagttttttagagtttaacatttgataaaaaa 3562
Db 77 GGCAGTAACTGCTACTACAATACTACTGAGTGTGTTGTGAGAGTTAAACATTTGATAATAAA 18

QY 3563 ctgacctgtttaacttc 3579
Db 17 CTGCGCTGTTTAACTTC 1

RESULT 13
AX093263 396 bp DNA linear PAT 30-MAR-2001
LOCUS
DEFINITION Sequence 81 from Patent WO0118046.
ACCESSION AX093263
VERSION AX093263.1 GI:13509712
```

```
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Xu,J. and Stolk,J.A.
Ovarian tumor sequences and methods of use therefor
JOURNAL Patent: WO 0118046-A 81 15-MAR-2001;
CORIXA CORPORATION (US)
FEATURES
SOURCE Location/Qualifiers
1. .396
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1. .396
/notes="n = A,T,C or G"
BASE COUNT 101 a 96 c 88 g 101 t 10 others
ORIGIN

Query Match 9.6%; Score 343.8; DB 6; Length 396;
Best Local Similarity 93.7%; Pred. No. 7.1e-72;
Matches 374; Conservative 0; Mismatches 22; Indels 3; Gaps 2;

QY 1186 gcagctgaagttcagcaggtgctgaatcgattctcctcgccctctcattccacttcca 1245
Db 1 GCAGCTGAAGTTTCAGCAGGTGCTGAATCGATTCTCTCGGCCCTCTCATTTCCACTTCCA 60
QY 1246 accctccattattccagttactacctcagcaattttgtgccccctacaaatgttagagac 1305
Db 61 ACCCTCCCATTTATTCAGTACTACCTCAGCAATTTTGTGCCCTTACAAATGTTAGAGAC 120
QY 1306 tgtatacgcttcagagctcttccctatgtagccacacattgagacatccctggatttcctg 1365
Db 121 TGTATAGCCCTTCAGGTCTTCCTATGCAGCCCAATTTGAGGACATCTGCACTTTCTGTG 180
QY 1366 gggaggtccccacagatattctactcatggggttcacatgggtttgaatcaccagggc 1425
Db 181 GGGAGGTTCCGCCACAGATATTCGTACTCATGGGTTTCACATGGTTTGAATCACCAGGN 240
QY 1426 cgcccatcaggagatgctttatccagatgaagctctcgacagagacattatggctgca 1485
Db 241 CGCCCATCAGGAGATGCTTTATCCAGATGAAGTCTCGGACAGANCAATTTATGGCTGCA 300
QY 1486 cagaagtgtcataaaaaacatgaagcacagatatgttgaagctcttcagtgctcagct 1545
Db 301 CAGAGTGGCATATAAAAAACATGAAGCACAGATATGTTGAAG--TTTTCAGTGTCACT 358
QY 1546 gaggagatgaacttgtgttaatgggggggcactttaaat 1584
Db 359 GANGANA-GAACATTGNNGTANNNGGGGNACTTTAAAT 396

RESULT 14
AX193162 329 bp DNA linear PAT 15-AUG-2001
LOCUS
DEFINITION Sequence 729 from Patent WO0149716.
ACCESSION AX193162
VERSION AX193162.1 GI:15211113
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J.,
Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.
Compounds for immunotherapy and diagnosis of colon cancer and
methods for their use
JOURNAL Patent: WO 0149716-A 729 12-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
SOURCE Location/Qualifiers
1. .329
```


misc_feature 69663..85425
/note="assembly_fragment"
misc_feature 85526..127682
/note="assembly_fragment"
misc_feature 127783..146089
/note="assembly_fragment"
clone_end:T7

BASE COUNT 40241 a 30457 c 30839 g 43752 t 800 others
ORIGIN

Query Match 8.5%; Score 303.4; DB 2; Length 146089;
Best Local Similarity 99.7%; Pred. No. 4.8e-62;
Matches 304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 901 qgtacttccaatgaggtagccagtttctctccaagaaaatcaagtcattgttcgcatg 960
Db 118645 GGTACTTCCAATGAGGTAGCCAGTTTCTCTCCAAGGAAAATCAAGTCATTGTCCGCATG 118704
QY 961 cgggggctccctttcacggccacagctgaagaagtgtggccctctttggacagcattgc 1020
Db 118705 CGGGGGCTCCCTTTCACGGCCACAGCTGAAGAGTGGTGGCCTTCTTTGGACACGCAATGC 118764
QY 1021 cctattactgggggaaagggaagcattcctctttgtcacctaccagatggttaggccaaca 1080
Db 118765 CCTATTACTGGGGGAAAGGAAGGCATCCTCTTTGTCACTACCCAGATGGTAGGCCAACA 118824
QY 1081 ggggacgttttgcctctttgctgtgaggaatatcacagaatgcgttgagggaagcat 1140
Db 118825 GGGGACGCTTTTGTCTCTTTCCTGTGGGAAATATCACAGAAATGCGTTGAGGAAGCAT 118884
QY 1141 aaagacttgttggttaaaagatacatgtgaactcttcaggagcacagcagctgaagttcag 1200
Db 118885 AAAGACTTGTGGTAAAGATACATTGAACCTTTCAGGAGCACAGCAGCTGAAGTTTCAG 118944
QY 1201 caggt 1205
Db 118945 CAGGT 118949

Search completed: August 15, 2002, 14:11:07
Job time: 30683 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 04:42:29 ; Search time 8772.8 seconds
(without alignments)
5515.523 Million cell updates/sec

Title: US-09-697-206A-1

Perfect score: 3585

Sequence: 1 ctttttggatcactgctgg.....gctgtttaatctcaaaaa 3585

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpi:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1062.8	29.6	1137	9	AL545672
c 2	986.8	27.5	1038	9	AL545648
3	769.6	21.5	777	10	BG182742
4	762.2	21.3	791	10	BG191000
5	741.6	20.7	786	10	BG216627
6	715.2	19.9	798	10	BG193725
7	666	18.6	768	10	BG209741
8	653.2	18.2	736	10	BG286508
9	637	17.8	782	10	BG217686
10	633.8	17.7	936	10	BG291485
11	629.6	17.6	913	10	BI557167
12	627.6	17.5	960	10	BG287429
13	619.8	17.3	819	10	BG184305
14	618	17.2	904	10	BE870387
15	615.8	17.2	890	10	EG260588
16	609.6	17.0	801	10	BI555187
c 17	608.4	17.0	819	10	BG218557

18	604.4	16.9	644	10	BG195224
19	603.4	16.8	646	10	BG217687
20	594	16.6	839	10	BG203925
21	593	16.5	804	10	BG188511
22	583	16.3	908	10	BI452459
c 23	581.2	16.2	628	10	BE710214
c 24	580.8	16.2	596	10	BF001941
c 25	560.6	15.6	729	10	BI694471
c 26	558	15.6	645	9	AW861027
c 27	557.6	15.6	667	9	AV698922
c 28	557.4	15.5	559	10	BM127059
c 29	553.6	15.4	594	9	AW363149
c 30	551.2	15.4	646	9	AW963406
c 31	550.2	15.3	574	9	AW382909
c 32	547.4	15.3	574	10	BG289507
c 33	545.2	15.2	650	10	BG403930
c 34	545	15.2	545	9	AI990751
c 35	544.8	15.2	548	9	AI923923
c 36	543	15.1	563	9	AI961474
c 37	537.4	15.0	542	9	AI813884
c 38	530.8	14.8	534	9	AA843844
c 39	530.6	14.8	865	10	BI852872
c 40	525.8	14.7	530	9	BE219683
c 41	525.4	14.7	639	9	AV694407
c 42	524.4	14.6	528	10	BI963927
c 43	522.8	14.6	526	9	AI301132
c 44	522.4	14.6	525	9	AI690278
c 45	520.8	14.5	524	9	AI963119

ALIGNMENTS

RESULT 1

AL545672
LOCUS AL545672 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI009YD20 5
DEFINITION prime, mRNA sequence.
ACCESSION AL545672
VERSION AL545672.1 GI:12878154
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1137)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..1137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI009YD20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="vector: pcwvSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 279 a 297 c 266 g 287 t 8 others

ORIGIN

```

Query Match      29.6%; Score 1062.8; DB 9; Length 1137;
Best Local Similarity 98.0%; Pred. No. 9.7e-206;
Matches 1103; Conservative 8; Mismatches 1; Indels 14; Gaps 3;

Qy 753 tcagggctgaaggaacgagagctctgtgttagttgttaagtgaaggagaccagacct 812
Db 1 TCAGGGCTGAAGGAACGAGAGCTCTGGTTAGTTTGTAACTKAGGAGCACCAGACCT 60

Qy 813 agcactacagaggcacaaacatcatggtggaccggtatattgaggttttacaagaac 872
Db 61 AGCCTACTAGAGGACACAAACATCATGAGGACCCGCTATATGAGGTTTACAAGCAAC 120

Qy 873 aggtgaagattcttctaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 932
Db 121 AGGTGAAGATTCTCTAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180

Qy 933 caagaaatacaagctattgttcgcatgagggggtcccttccacgagctgaaga 992
Db 181 CAAGAAATCAAGTCATG-TGCGATGCGGGGCTCCCTTTCACGGCCACAGCTGAAGA 239

Qy 993 agtgggtccttcttgagacagattcccttattactggtggggaaggaagcattcttt 1052
Db 240 AGTGGTGGCTTCTTGGACAGCATTCCTTACTTACTGGGGGAAAGMAGGCATCCCTTT 299

Qy 1053 tgcacctaccagatggttaggccaacagggagcgttttcttcttctgtgtgagga 1112
Db 300 TGTCACCTACCCAGATGGTAGGCCAACAGGGAGCGCTTTTTCCTCTTGGCTGTKAGGA 359

Qy 1113 atatgcagaatcgttgagggaagcataaagactgttggtgtaaaagatacatgaact 1172
Db 360 ATATGCAGAAATCGTTGAGGAAGCATAAAGACTTGTGGGTAAABATACATGAAT 419

Qy 1173 ctacagagacagcagctgaagttgagcaggtgtctgaatcattctcctggccctct 1232
Db 420 CTTCAGGAGCACAGCAGTGAAGTTTCAGCAGGTGCTGAATCGATTTCTTCCTGGGCCCTCT 479

Qy 1233 cattccacttccaccctccatttccagttactactaccctcagcaatttggccccctac 1292
Db 480 CATTCACCTTCCAAACCCCTCCCATTTATCCAGTACTACCTCAGCAGMAATTTGGCCCCCTAC 539

Qy 1293 aaatgttagactgtatagcctctcaggtcttccctatgcagcccaaatgaggacat 1352
Db 540 AAATGTTTAGACTGTATACGCCCTTCGAGGTCTTCCCTATGCAGCCACAATGAGGACAT 599

Qy 1353 cctgatttctggggaggtcgcacagatattcgtactcaggtgttcaatcaggtttt 1412
Db 600 CCTGGATTTCTTGGGGAGTTGCCACAGATATTCGTACTCATGGGGTTCACATGGTTTT 659

Qy 1413 gaatcacaggccgcccacatcagagatgccttttccagatgaagtctgcggacagc 1472
Db 660 GAATCACAGGGCCGCCCATCAGGAGATGCTTTTATCCAGATGAAGTCTCGGACAGAGC 719

Qy 1473 attatggctgcagagaagtgtctataaaacacatgaaggacagatgtgtgaagcttt 1532
Db 720 ATTTATGGCTGCACAGAAGTGTCTATAAAAAACATGAAGACAGATATGTGTGAAGTCTT 779

Qy 1533 tcagtgttcagctagcagatgaacttgttgaatgggggacatttaaatcgaatatgg 1592
Db 780 TCAGTGTTCAGCTAGGAGATGAATTTGTGTTAATGGGGGACATTTAAATCGAAATGG 839

Qy 1593 ctatccccccagc-----ccatgcctgtctcctccctccctacacatttccagc 1640
Db 840 CTTATCCCCACCGCCATGTAAAGTTTACCATGCTGTCTCTCTCCCTTACATTTCCAGC 899

Qy 1641 tccgtgtcagttatttccacagaagctgcccatttaccagccctctgtgatttgaatcc 1700
Db 900 TCCGTGCTGCAATTTATCTTACAGAAGTGCCTATTTACAGCCCTCTGTGTGATTTTGAATCC 959

Qy 1701 acgagcactgcagccctccacagcgtactaccagcagggcactcagctcttctgaacta 1760
Db 960 ACGAGACTGCAGGCCCTCCACAGCGTACTACCCAGCAGGCACTCAGCTCTTCATGAATTA 1019

```

```

Qy 1761 cacagcgtactatccccagccccccaggttcgcttaatagtcttggctacttccctacagc 1820
Db 1020 CACAGCGTACTATCCAGAGCCCCCAGGTTGCGCTAAATAGTCTTGGCTACTTCCCTACAGC 1079

Qy 1821 tgctaatcttagcgtgtcctccacagcctggcacggtgggtcaga 1866
Db 1080 TGCTAAATCTTAGCGGTGTCCTT-CACAGCCTGGCAGCGTGGTCTAGA 1124

RESULT 2
AL545648/c
LOCUS AL545648 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI009YD20 3
DEFINITION prime, mRNA sequence.
ACCESSION AL545648
VERSION AL545648.1 GI:12878130
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1038)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..1038
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI009YD20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 322 a 173 c 164 g 351 t 28 others
ORIGIN

Query Match      27.5%; Score 986.8; DB 9; Length 1038;
Best Local Similarity 96.3%; Pred. No. 2.7e-190;
Matches 993; Conservative 24; Mismatches 13; Indels 1; Gaps 1;

Qy 2446 ttccctgtcttgcataatcatgcctcttctgtaagtaattcgaagataagatcttgg 2505
Db 1030 TTCCCTGTCTCTGTCATATCATGCTCTTCTGCTAGTAATTCAAGCATAAGATCTKGG 971

Qy 2506 aataataaaatcaacaatcttaggagaaagaataaaattgtatttcccagctcttggc 2565
Db 970 AATAATAAAATCAAAATCTTAGGAGAAAGTTTAAATTTGTTATTTCCACGCTCTCTGGC 911

Qy 2566 catgatgatctctatgattaaaaacaaattataatttttaaacacctgaagataaattag 2625
Db 910 CATGATGATCTTATGATTTAAAAACAAATTAATTTTAAACACCTGAAGATATATTAG 851

Qy 2626 aagaaattgtgcacccctccacaaacatacaaaagtgttaaaagtgttgatctttctcag 2685
Db 850 AAGAAATTTGTGCACCCCTCCCAACAAACATACAAAGTTTAAAGTTTGGATCTTTTCTCAG 791

Qy 2686 caggtatcagttgtataataatgaattagggcccaaatgcaaaacgaataaagaagcagc 2745
Db 790 CAGGTATCAGTTGTAAATAATGAATTAGGGGCCAAAATGCAAAACGAAAAATGAAGCAGC 731

```

Qy 2746 tacatgtagttagtaatttcttagttgaactgttaattgaattatgttggtcttcacatgtat 2805
|||||
Db 730 TACATGTAGTTAGTAATTCTTAGTTGAACGTGAATTGAATATTGTGGCTTCATATGTAT 671

Qy 2806 tatittatattgtacttttttccattatattgagtggttgactttaataagagaaattccat 2865
|||||
Db 670 TATTATATATTGTACTTTTTTTTCATTATTGATGGTTTGGACTTTAATAAGAGAAATTCAT 611

Qy 2866 agtttttaattatccagagtgagacaatttgaacagtgattcttagaaaaacaatacacat 2925
|||||
Db 610 AGTKTTTAATATCCGAGAAGTGAGCAATTTGAACAGGTGATTCFAGAAAAACAATACACAT 551

Qy 2926 aactgaacagagtgaaatgccttatatatattatgatagcctttaaacctttttccctccta 2985
|||||
Db 550 AACTGAACAGAGAAGTGAAGACTBBATATATATATGATAGCCTTAAACMATTTTCCCTCWAAT 491

Qy 2986 gcccttaactgcacaaataattaaaccttttaagacataggaactatagtcagcatgctaga 3045
|||||
Db 490 GCCTTAACTGACAAATAATTAACTTKTAAAGCANTAGGACTATAGTCAGCATGCTAGA 431

Qy 3046 ctgagagtgaaacactgatgcaatttagaacacaggtactgctgctcagtggttttaacacta 3105
|||||
Db 430 CTGAGAGTAAACACTGATGCAATTAGAACAGGTACTGATGCTGACGTGKWTACACTA 371

Qy 3106 tgtttagctgtgtttatgcctataaaagtgcgaattattagacactagctagctagctgct 3165
|||||
Db 370 TGTNTAGCNGTGTATTATGCTATATAAARNVCAATATTAGACACTAGTAGTACSGCBGCSB 311

Qy 3166 catgtaactccaaagaaaaacaggatttcatttaagtgcattgaattgagatatttctctaa 3225
|||||
Db 310 CATGTAACBCCAAAGAAAACAGGATKTCATTAAAGTGCAATKGAATGTGCBATBTCTCAA 251

Qy 3226 gttaactaatgttccttgccttgaatgcaatgcagtcgctgagatttatagagctcatttt 3285
|||||
Db 250 GTBACTCATATGTGCTTGTGKBNAATGCAATGCCGTGCACATTATATGTGCTGCTATT 191

Qy 3286 ttattttctgtcattactttaacaccccttaaaaggagagagaaacattctcttcttcagc 3345
|||||
Db 190 TTAATTTTC-GTGCATTACTTTAAACACCTTAAAGGAGAAGCAAAACATTTCCTTCTTCAGC 132

Qy 3346 tgactggcaatggccctttaactgcaataggaagaaaaaagggttgtgtgaaaaat 3405
|||||
Db 131 TGACTGGCAATGGCCCTTTAACTGCAATFAGGAAGAAAAAAGGTTTGTGTGAAAAAT 72

Qy 3406 tgggtgaactggcacttaagatcgaaagaaattctgtacttgatgctctaaagatg 3465
|||||
Db 71 TGGTGATAAATGGGCATTAAAGATCGAAAGAAATTTCTGTATCTTGTATGCTTAAGATG 12

Qy 3466 cccaaagctgc 3476
|||||
Db 11 CCCAAAGCTGM 1

RESULT 3
Bg182742 777 bp mRNA linear EST 21-APR-2001
LOCUS Bg182742
DEFINITION RST1619 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION Bg182742
VERSION Bg182742.1 GI:13704429
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 777)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,
E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,
J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression

JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc. Cleveland, OH 44115, USA
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 552.
FEATURES
Location/Qualifiers
source 1..777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_lines="HT1080"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
BASE COUNT 251 a 139 c 136 g 251 t
ORIGIN

Query Match 21.5%; Score 769.6; DB 10; Length 777;
Best Local Similarity 99.5%; Pred. No. 3.8e-146;
Matches 772; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2038 agtgttgaaagatgatatggtgatcttgaaacctccagacacagaaacctctctagcaaa 2097
Db 2 AGTGTGTTGAAAGATGATGGTGATCTTGAACCTCCAGACACAGAAACATTTCTAGCAAA 61

Qy 2098 ttcagggaagtgttctacactcaggtcgagatttttcagcaactgattgacaaa 2157
Db 62 TTCAGGGGAAGTTGCTACACTCAGGCTGCAGTATTTTCACCAACTTGATGGACAAA 121

Qy 2158 cgggcctgtcccttattcttgggtgagtgaaataatttgagctagtgaaagccaaatcgt 2217
Db 122 CGGGCTGTGCTTATCTTTTGGCGAGTGAAAAAATTTGAACCTAGTGAAGCCAAATCGT 181

Qy 2218 aacttaacggaagcagcatgacataccctggctcttctgctgattgcaaataggcattta 2277
Db 182 AACTTACAGCAAGCAGCATGCGACATACCTGGCTCTTTGCTGATTGCCAAATAGGCATTTA 241

Qy 2278 aaatgtgaattgggaatcagatgtctccattacttccagttaaagtgagcatataggtgt 2337
Db 242 TAATGTGAATTTGGAAATCAGATGCTCTCCATTAATTCCAGTTACAGTGGCATTATAGGTGT 301

Qy 2338 ttcctaagttttaagctcttggtataaaaaactccaccagtgctaccatctccaccatgaac 2397
Db 302 TTCCTAAGTTTAAAGTCTTGGATAAAAACTCCACCAGTGTCTACCATCTCCACCATGAAC 361

Qy 2398 tctgttaagggaagcttcattttgttatattcccgctcttttcttcttcttccctgtctt 2457
Db 362 TCTGTTAAGGAAGCTTCATTTTGTATATTCCCGCTCTTTTCTCTCAATTTCCCTGTCTT 421

Qy 2458 ctgcataatcatgcctctcttgctaagtaattcaagcataagatctcttggaataataaac 2517
Db 422 CTGCATAATCATGCCCTTCTTGGCTAAGTAATTCACACATAAGATCTTGGAAATAAATATC 481

Qy 2518 acaatcttaggagaagaataaaattgttatttcccagctctcttggccatgatgatc 2577
Db 482 ACAATCTTAGGAGAAAGAAATAAAATTTGTTATTTCCAGTCTCTTGGCCATGATGATATC 541

Qy 2578 ttatgattaaaaacaaataaattttaaacaacctgaagataaaattagaagaaattgtgc 2637
Db 542 TTATGATTAAAAACAATAATTAATTTTAAACACCTGAAAGATAAAATTAGAGAAATTTGTC 601

Qy 2638 accctccacaaacatacaaaagttttaaagtttgagatcttttctcagcaggtatcagtt 2697
Db 602 ACCCTCCACAAAACATACAAAGTTTAAAGTTTGGATCTTTTCTCAGCAGGTATCAGTT 661

```
QY 2698 gtaataatgaattagggcccaaaatgcaaaacgaaagaaagcagctacatgtagtta 2757
|||||
Db 662 GTAAATAATGAATTAGGGGCCAAAATCAAAACGAAAATGAAGCAGCTACATGTAAGTTA 721

QY 2758 gtaattctagttgaactgtaattgaattgttggttcattatgtattttat 2813
|||||
Db 722 GTAAATTTCTAGTTGAAGTCTGAATTTGAATATTTGGCTTCATATGTAATTTAT 777

RESULT 4
BG191000
LOCUS RST10214 Athersys RAGE Library Homo sapiens cDNA, mRNA linear EST 21-APR-2001
DEFINITION BG191000
ACCESSION BG191000
VERSION BG191000.1 GI:13712815
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM human.
REFERENCE 1 (bases 1 to 791)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Levenchal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Kiika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scaine@atersys.com
High quality sequence stop: 547.
Location/Qualifiers
1..791
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 251 a 138 c 139 g 260 t 3 others
ORIGIN

Query Match 21.3%; Score 762.2; DB 10; Length 791;
Best Local Similarity 98.8%; Pred. No. 1.2e-144;
Matches 778; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 2038 agtgtttgaagatgtatgtgattgaaacctccagacacaaagaaactctagcaaa 2097
|||||
Db 3 AGTGTTTGAAAGATGATGGTGATCTTTGAAACCTCCAGACACAGAAACTCTAGCAAA 62

QY 2098 ttacggggaagtgtgtacacactcaggctgcagtgattttcagcaaaacttgattggacaaa 2157
|||||
Db 63 TTCAGGGGATGTTGTCTACACTCAGGCTGCAGTATTTTCAGCAAACTTGATTGGACAAA 122

QY 2158 cgggcctgtcccttatcttttggtagtgaaataattgagctagtgaaagcgaatacgt 2217
|||||
Db 123 CGGGCCCTGTGCCTTATCTTTTGGTGGAGTGAATAAATTTGAGTAGTGAAGCCAAATCGT 182

QY 2218 aacttacgaacagcagcatgcagcatcacctggctcttctgctgattgcaaataggcattta 2277
|||||
Db 183 AACTTACGCAAGCAGCATGCAGCATACCTTGGCTCTTGTGCTGATTGCAATAGGCATTTTA 242
```

```
RESULT 5
BG216627
LOCUS RST36320 Athersys RAGE Library Homo sapiens cDNA, mRNA linear EST 21-APR-2001
DEFINITION BG216627
ACCESSION BG216627
VERSION BG216627.1 GI:13742648
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 786)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Kiika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scaine@atersys.com
High quality sequence stop: 557.
Location/Qualifiers
1..786
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source
```

```

/clone_lib="Athersys RAGE Library"
/cell_line="Ht1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is Ht1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in Ht1080 under normal circumstances."
252 a 152 c 145 g 234 t
BASE COUNT
ORIGIN

```

Query Match	20.7%;	Score	741.6;	DB	10;	Length	786;
Best Local Similarity	98.8%;	Pred. No.	1.8e-140;				
Matches	747;	Conservative	0;	Mismatches	9;	Indels	0;
Gaps	0;						
Qy	1949	ttatacacaaaatgaccagccagagcactcaccacaagaatgggtttgtatttaaggcc	2008				
Db	30	TTATACACAAACGACCGCAGGACTACCCAAAGATGGTGTGTATTTAAGGCG	89				
Qy	2009	cccagcagttagaacatcctcagaaaaagaatgtttgaaaagtgtatggtgatctctgaaa	2068				
Db	90	CCAGCAGTTAGAGCATCCTCAGAAAAAGATGTTTGAAGATGATGGTGATCTTGAAA	149				
Qy	2069	cctccagacacaaagaaaactctcagcaaatcagggggaagtgttctcacactcagcgtgc	2128				
Db	150	CCTCCAGACACAAGAAAATCTCTAGCAAAATTCAGGGGAAGTTGTCTACACTCAGGCTGC	209				
Qy	2129	agtatcttcacgaaacttgatggacaaacgggcctgtgccttactctttgggtggagtga	2188				
Db	210	AGTATTTTCAGCAAACTTGATTTGGNCAAAACGGGCTGTGCCCTTATCTTTTGGTGGAGTGA	269				
Qy	2189	aaaaatttgagctagtgagcgcacaaatcgttaacttacagcagcagcatgcagcataccctg	2248				
Db	270	AAAAATTTGAGCTAGTGAAGCAAAATCATAACTTACAGCAAGCAGCATGCAGCATACCTG	329				
Qy	2249	gctcttctcgtatgcacaaataggcattttaaagtgaattggaaatcagaatgtctccatt	2308				
Db	330	GCTCTTCTGCTGATGCAAAATAGGCATTTAAATGNGAATTTGGAATCAGATGTCFCCATT	389				
Qy	2309	acttccagttaaagtgccatcatagtggttctcctaagttttaaagctcttggaataaaactc	2368				
Db	390	ACTTCCAGTTAAAGTGGCATCATAGGTCTTCTTAAAGTTTAAAGTCTTTGGATAATACTC	449				
Qy	2369	cacagtgctcacatctccaccatgaactctgttaaaggaagcttcattttgttatctc	2428				
Db	450	CACCAGTGTCTACCATCTCCACCATGAATCTTGTTTAAGGAAGCTTCATTTTGTATATTC	509				
Qy	2429	ccgctctttctcttcatttccctgctctctgcataatcagctctcttgctaagtaatt	2488				
Db	510	CGGCTCTTTTCTCTTCATCTCCCTGCTCTCTGCAATAATCATGCCCTCTTGCTAAGTAATT	569				
Qy	2489	caagcataagatcttggaaataaaaaatcaaatctttaggagaaagaataaaattgttat	2548				
Db	570	CAAGCATAAAGATCTTGGAATAATAAAATCACAATCTTAGGAGAGAGATAAAATTTGTTAT	629				
Qy	2549	tttccagctctctggccatgatgatcttatgattaaaaacaaatataatttttaaac	2608				
Db	630	TTTCCCAGTCTCTTGGCCATGATGATATCTATGATTTAAAAACAATTTTAAATTTTAAAAAC	689				
Qy	2609	acctgaagataaaattagaagaaattgtgcacctccacaaacatacaaaagttttaaagt	2668				
Db	690	ACCAGAGATAAATTAGAAGAAATTTGTCACCCCTCCACAAACATACAAAGTTTAAAAAGT	749				
Qy	2669	ttgatcttttctcagcaggtatcagttgtaata	2704				
Db	750	TTGGATCTTTTCTCAGCAGGTATCAGTTGTAAAAA	785				

RESULT	6
BG193725	
LOCUS	BG193725
DEFINITION	RST12861 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
	798 bp linear EST 21-APR-2001

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE	JOURNAL	MEDLINE	COMMENT
-------	---------	---------	---------

FEATURES	SOURCE
1. The first two columns are labeled "FEATURES" and "SOURCE".	
2. The third column is labeled "FEATURES".	
3. The fourth column is labeled "SOURCE".	
4. The fifth column is labeled "FEATURES".	
5. The sixth column is labeled "SOURCE".	
6. The seventh column is labeled "FEATURES".	
7. The eighth column is labeled "SOURCE".	
8. The ninth column is labeled "FEATURES".	
9. The tenth column is labeled "SOURCE".	
10. The eleventh column is labeled "FEATURES".	
11. The twelfth column is labeled "SOURCE".	
12. The thirteenth column is labeled "FEATURES".	
13. The fourteenth column is labeled "SOURCE".	
14. The fifteenth column is labeled "FEATURES".	
15. The sixteenth column is labeled "SOURCE".	
16. The seventeenth column is labeled "FEATURES".	
17. The eighteenth column is labeled "SOURCE".	
18. The nineteenth column is labeled "FEATURES".	
19. The twentieth column is labeled "SOURCE".	
20. The twenty-first column is labeled "FEATURES".	
21. The twenty-second column is labeled "SOURCE".	
22. The twenty-third column is labeled "FEATURES".	
23. The twenty-fourth column is labeled "SOURCE".	
24. The twenty-fifth column is labeled "FEATURES".	
25. The twenty-sixth column is labeled "SOURCE".	
26. The twenty-seventh column is labeled "FEATURES".	
27. The twenty-eighth column is labeled "SOURCE".	
28. The twenty-ninth column is labeled "FEATURES".	
29. The thirtieth column is labeled "SOURCE".	
30. The thirty-first column is labeled "FEATURES".	
31. The thirty-second column is labeled "SOURCE".	
32. The thirty-third column is labeled "FEATURES".	
33. The thirty-fourth column is labeled "SOURCE".	
34. The thirty-fifth column is labeled "FEATURES".	
35. The thirty-sixth column is labeled "SOURCE".	
36. The thirty-seventh column is labeled "FEATURES".	
37. The thirty-eighth column is labeled "SOURCE".	
38. The thirty-ninth column is labeled "FEATURES".	
39. The fortieth column is labeled "SOURCE".	
40. The forty-first column is labeled "FEATURES".	
41. The forty-second column is labeled "SOURCE".	
42. The forty-third column is labeled "FEATURES".	
43. The forty-fourth column is labeled "SOURCE".	
44. The forty-fifth column is labeled "FEATURES".	
45. The forty-sixth column is labeled "SOURCE".	
46. The forty-seventh column is labeled "FEATURES".	
47. The forty-eighth column is labeled "SOURCE".	
48. The forty-ninth column is labeled "FEATURES".	
49. The fiftieth column is labeled "SOURCE".	
50. The fifty-first column is labeled "FEATURES".	
51. The fifty-second column is labeled "SOURCE".	
52. The fifty-third column is labeled "FEATURES".	
53. The fifty-fourth column is labeled "SOURCE".	
54. The fifty-fifth column is labeled "FEATURES".	
55. The fifty-sixth column is labeled "SOURCE".	
56. The fifty-seventh column is labeled "FEATURES".	
57. The fifty-eighth column is labeled "SOURCE".	
58. The fifty-ninth column is labeled "FEATURES".	
59. The sixtieth column is labeled "SOURCE".	
60. The sixty-first column is labeled "FEATURES".	
61. The sixty-second column is labeled "SOURCE".	
62. The sixty-third column is labeled "FEATURES".	
63. The sixty-fourth column is labeled "SOURCE".	
64. The sixty-fifth column is labeled "FEATURES".	
65. The sixty-sixth column is labeled "SOURCE".	
66. The sixty-seventh column is labeled "FEATURES".	
67. The sixty-eighth column is labeled "SOURCE".	
68. The sixty-ninth column is labeled "FEATURES".	
69. The seventieth column is labeled "SOURCE".	
70. The seventy-first column is labeled "FEATURES".	
71. The seventy-second column is labeled "SOURCE".	
72. The seventy-third column is labeled "FEATURES".	
73. The seventy-fourth column is labeled "SOURCE".	
74. The seventy-fifth column is labeled "FEATURES".	
75. The seventy-sixth column is labeled "SOURCE".	
76. The seventy-seventh column is labeled "FEATURES".	
77. The seventy-eighth column is labeled "SOURCE".	
78. The seventy-ninth column is labeled "FEATURES".	
79. The eightieth column is labeled "SOURCE".	
80. The eighty-first column is labeled "FEATURES".	
81. The eighty-second column is labeled "SOURCE".	
82. The eighty-third column is labeled "FEATURES".	
83. The eighty-fourth column is labeled "SOURCE".	
84. The eighty-fifth column is labeled "FEATURES".	
85. The eighty-sixth column is labeled "SOURCE".	
86. The eighty-seventh column is labeled "FEATURES".	
87. The eighty-eighth column is labeled "SOURCE".	
88. The eighty-ninth column is labeled "FEATURES".	
89. The ninetieth column is labeled "SOURCE".	
90. The ninety-first column is labeled "FEATURES".	
91. The ninety-second column is labeled "SOURCE".	
92. The ninety-third column is labeled "FEATURES".	
93. The ninety-fourth column is labeled "SOURCE".	
94. The ninety-fifth column is labeled "FEATURES".	
95. The ninety-sixth column is labeled "SOURCE".	
96. The ninety-seventh column is labeled "FEATURES".	
97. The ninety-eighth column is labeled "SOURCE".	
98. The ninety-ninth column is labeled "FEATURES".	
99. The hundredth column is labeled "SOURCE".	

BASE COUNT
ORIGIN

Query Ma	Best Loc	Matches
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1
10	1	1
11	1	1
12	1	1
13	1	1
14	1	1
15	1	1
16	1	1
17	1	1
18	1	1
19	1	1
20	1	1
21	1	1
22	1	1
23	1	1
24	1	1
25	1	1
26	1	1
27	1	1
28	1	1
29	1	1
30	1	1
31	1	1
32	1	1
33	1	1
34	1	1
35	1	1
36	1	1
37	1	1
38	1	1
39	1	1
40	1	1
41	1	1
42	1	1
43	1	1
44	1	1
45	1	1
46	1	1
47	1	1
48	1	1
49	1	1
50	1	1
51	1	1
52	1	1
53	1	1
54	1	1
55	1	1
56	1	1
57	1	1
58	1	1
59	1	1
60	1	1
61	1	1
62	1	1
63	1	1
64	1	1
65	1	1
66	1	1
67	1	1
68	1	1
69	1	1
70	1	1
71	1	1
72	1	1
73	1	1
74	1	1
75	1	1
76	1	1
77	1	1
78	1	1
79	1	1
80	1	1
81	1	1
82	1	1
83	1	1
84	1	1
85	1	1
86	1	1
87	1	1
88	1	1
89	1	1
90	1	1
91	1	1
92	1	1
93	1	1
94	1	1
95	1	1
96	1	1
97	1	1
98	1	1
99	1	1
100	1	1

Qy	2043
Dp	8

Qy	2103
Db	68

Qy	2163
Db	128

Qy	2223
Db	188

QY	2283
Db	248

Qy	2343
Db	308

QY	2403
Db	368

QY 2463

BG193725
BG193725.1 GI:13715412
EST.
human.

1 (bases 1 to 798)
Harrington, J.J., Sherf,
Mammalia; Eutheria; Pri
Eukaryota; Metazoa; Cho

Creation of genome-wide
activation of gene expression
Nat. Biotechnol. 19 (5)
21227151
Contact: Scott J. Cain

Location/Quali
1. .798

257 a 140 c 1

ch 19.9%; S
l Similarity 97.2%; P
770: Conservative 0:

ttgaagatgtatggtgatcttgaa
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
TACAAAGAGTATGCTTCTTGAA

gggaagttgtctacactcaggctg
|||||
GGGAAGTTGTCTACACTCAGGCTG

tgtagccttatcttttggtggagtg
|||
AGTGCCCTATCTTTGGTGGAGTG

cagcaagcagcatgcagcatacct
|||||
CAGCAAGCAGCATGCAGCATACCT

ggaatttgaatcagatgtctccat
|||||
GGAATTGGAATCAGATGTCTCCAT

aggtttaagctctggataaaaact
|||||
AGTTTAAGCTCTGGATAAAAAC

aaaggaagcttcatttttgatatatt
|||||
AAGGAAGCTTCATTTTGTATATT

aatcatgcccttcttgctaagtaaat

BG193725
 BG193725.1 GI:13715412
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 798)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
 , Lerner,L., Costanzo,D., Mcelligott,K., Boozer,S., Mays,R., Smith
 ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
 J., Danzig,J. and Ducar,M.
 Creation of genome-wide protein expression libraries using random
 activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)
 21227151
 Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 548.
 Location/Qualifiers
 1..798
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Athersys RAGE Library"
 /cell_line="HT1080"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."

[illegible]

```

|||||
Db 428 TAATCATGCTCTTCTGCTAAAGTAATTCAGCATAAGATCTTTGGAAATAATAAAATCACAAAT 487
Qy 2523 cttaggagaagaataaaattgtattttccagctctcttgccatgatgatattttatg 2582
Db 498 CTTAGAGAAAGAAATAAAATGTTATTTTCCAGTCTCTTGCCATGATGATATCTTATG 547
Qy 2583 attaaacaaattaaa-ttttaaacacctgaagataaaattagaagaattgtgcacc 2641
Db 548 ATTAACAAATTAATTTTAAACACCTGAAGATAAAATTAGAAGAAATTTGTGCACCC 607
Qy 2642 tccacaacatacaaaagttaaagtgttgatctttttctcagcaggtatcagttgtaa 2701
Db 608 TCCACAAACATACAAAGTTTAAAGTTTGGATCTTTTCTCAGCAGGTATCAGTTGTAA 667
Qy 2702 ataagtaattagggcccaaaatgcacaaacgaacaaatgaagcagctcacatgattagtaa 2761
Db 668 ATAAATGAATAGGGGGCAAAATGCCAAGC-AAATGAAGCAGCTACATGTAGTTAGTAA 726
Qy 2762 -ttctagttgaactgaattgaattatt-gtggcttcataatgtatttttatattgta 2819
Db 727 TTTTCTAGTTGAAGTGAATTAATTTGGGGCTTCATATGAATTAATTTTATATCGGA 786
Qy 2820 cttttttcaata 2831
Db 787 CTTTTTATTTA 798

```

```

RESULT 7
BG209741 768 bp mRNA linear EST 21-APR-2001
LOCUS R529267 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG209741
VERSION BG209741.1 GI:13731428
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Harrington,J.J., Sherr,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
E., Veloso,N., Kikla,A., Hess,J., Cothren,K., Lo,K., Offenbacher
J., Danzig,J., and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
2127151
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scaine@atersys.com
High quality sequence stop: 461.
Location/Qualifiers
1..768
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="H1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is H1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in H1080 under normal circumstances."
247 a 151 c 141 g 229 t

```

```

FEATURES
source
1..768
BG286508 736 bp mRNA linear EST 21-FEB-2001
LOCUS 602382504F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4500121 5',
DEFINITION mRNA sequence.
ACCESSION BG286508
VERSION BG286508.1 GI:13039446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 736)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Query Match 18.6%; Score 666; DB 10; Length 768;

```

```

Best Local Similarity 96.0%; Pred. No. 4.3e-125;
Matches 737; Conservative 0; Mismatches 25; Indels 6; Gaps 5;
Qy 1940 aggatggacttatcacacaaatgaccaggccagactcttaccacaaagaattggttga 1999
Db 1 AGGATGGACTTATACACAAATGACCAGGCCAGGACTCTACCCAAAGAAATGGTGTGTA 60
Qy 2000 ttttaaggcccccagcagtttagaacaactctcagaagaagtggttgaaagatgtatgg 2059
Db 61 TTTAAGGCCCCAGCAGTGTAGAACATCTCGAAGAAAGTGTGTTGAAAGATGTATGGTG 120
Qy 2060 atcttgaaacctccagacacacagaacaaactctcagaaatcagggggaagtttgtctaac 2119
Db 121 ATCTTTGAACCTTCCAGACACACAAACCTTCTAGCAAAATTCAGGGGAAGTTTGTCTAC 180
Qy 2120 tcaggtcagttatcttcagcaaaacttgatggcaaaacggcctgtgcttattttt 2179
Db 181 TCAGGCTCAGAAATTTTCAGCAAACTTGATTGGACAAACGGGCTGTGCTTATCTTTTG 240
Qy 2180 gtggagtgaaaaaatttgagctagtgaaagccaaaatcgtaacttacagcaagcagcatga 2239
Db 241 GTGGAGTGAAGAAATTTGAGCTAGTGAAGCCAAATCGTAACTTACAGCAAGCAGCATGCA 300
Qy 2240 gcataacctggctcttgctgattgcaaataggcatttaaaatgtgaatttggaatcagat 2299
Db 301 GCATACCTGGCTCTTGTGATTGCAAAATAGCAATTAATAATGTGAATTTGGAATCAGAT 360
Qy 2300 gtctccattacttccagttaaagtggcatcataggtgttctcctaagtttaagttcttga 2359
Db 361 GTCTCCATTACTTCCAGTTAAAGTGGCATCATAGGTGTTTCTTAAGTTTAAAGTCTTGA 420
Qy 2360 taaaaactccacagctgtctaccatctccacatgaactctgtaaggaagcttcaattt 2419
Db 421 TAAAAAATCCACACAGTGTCTACCATCTCCACCATGAACCTCTGTTAAGGAAGCTTCAATTT 480
Qy 2420 tqtatctccgctctttctctctcatttccctctctctctgcataatca-tgcctctctg 2478
Db 481 TGTATATTCCTCTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Qy 2479 ctaagtaattcaagcataagatc-ttggaataataaaaaacacaaatctttagggagaagaat 2537
Db 541 CTAAGTAATTCAAGCATAAGATCTTTGGAAATAATAAAATCAATTTCTTAGGAGAAAGAAAT 600
Qy 2538 aaaaattgtattttccagctctcttgccatgatgatcttattgattaaaaacaattaa 2597
Db 601 AAAATGTGTTATTTTCCAGTCTTTTGGGCTGTATGATATCTTATGATTAAAAACAGATTA 660
Qy 2598 aa-ttttaaaacacctgaagataaaattagaagaaa-ttgfcacacctccacaaacatac 2655
Db 661 AATTTTAAAAACACTGGAAACATAAATAAGAACAAATTTGGGCACCTCCACCAACATAC 720
Qy 2656 aaagttaaagaatt--tggatctttttctcagcaggtatcagttgtaa 2701
Db 721 AAAGGTTTAAAGTCCGGATCCTTTTCTCAACAGGAATCAGATGGAA 768

```


Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10365 row: g column: 02
High quality sequence stop: 730.
Location/Qualifiers
1. .736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4500121"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
202 a 189 c 159 g 186 t

BASE COUNT
ORIGIN

Query Match 18.2%; Score 653.2; DB 10; Length 736;
Best Local Similarity 97.4%; Pred. No. 1.7e-122;
Matches 721; Conservative 0; Mismatches 3; Indels 16; Gaps 5;
QY 1518 atattgtgaagcttttcagttcagctgagggatgaactttgttaattggtggggcagc 1577
Db 1 ATATGTTGAAGCTCTTCAG-GTTCAGCTGAGGATGAACGTGTGTATATGGGGGCAC 59
QY 1578 tttaaatcgaatggcttatcccccagc-----ccatgcctgtctctccctc 1625
Db 60 TTTAAATCGAAATGGCTTATCCCGCCGATGTAAGTTACCATGCTGTCTCTCCCTC 119
QY 1626 ctacacattccagctcctcctcagttatctctacagaagctgcattaccagcctc 1685
Db 120 CTACATATTTCCAGCTCCTGCTGCGATTTATCCFACAGAAGCTGCCATTTACCGCCCTC 179
QY 1686 tgtgatttgaatccagcagcactgcagcctccacagcgtactaccagcagcactca 1745
Db 180 TGTGATTTTGAATCCAGAGCACTGCAGCCCTCCACAGCTACTACCCAGGCACTCA 239
QY 1746 gctctcatgaactacacagcgtactatcccagccccccaggttcgcctaatgcttgg 1805
Db 240 GCTCTTCATGAATPACACAGCGTACTATCCAGCCCCCAGGTTTCGCTAATAGTCTGG 299
QY 1806 ctacttccctacagctgaactctagcgtgtctctccacagcctgcacggtggtcag 1865
Db 300 CTACTTCCCTACAGCTGCTAATCTTAGCGGTGTCCTCCACAGCTGGCAGGTTGGTCAG 359
QY 1866 aatgcaggcctggcctacatactgaggttaaggaatcttaacttcttccaaagtta 1925
Db 360 AATGCAGGCGCTGSCCTACAACTACTGGAGTTAAGGAATTTCTTAACCTTCTCCAGGTTA 419
QY 1926 ccagtatgcacccagagatgaacttatcacacacaaatgaccagccaggactctaccocaa 1985
Db 420 CCAGTATGCACCCAGAGATGGACTTATACACACAAATGACAGCCAGGACTCTTACCCAA 479
QY 1986 agaaggtttgtatttaagggcccccagcaggttagaagcattcctcagaagaagtggttg 2045
Db 480 AGAATGGG-TTGTATTTAAGGGCCCCAGCAGTATAGAACATCTCTCAGAAAAAGAGTGTGG 538
QY 2046 aaagatgatggtgatcttgaacctccagacacaaagaaactctcagcaaatcagggg 2105
Db 539 AAAGATGATGGTGATCTTGAACCTCCAGACACAAAGAAATCTTAGCAAAATTCAGGGG 598
QY 2106 aagttgtctacactcaggtgcagtagtatttcagcaactgatttgacaaagggcgtg 2165

Db 599 AAGTTTGTCTACACTCAGGCTGCAGTAGTATTTTCAGCAAACTTGATTGGACAAACGGGCGTG 658
QY 2166 tgcctattcttggtagtgagtgaaattgaactgaactgaactgaactgaactgaactga 2225
Db 659 TGCCTTATCTTTTGGTGAGTGAAAAAATTTGACCCAGTGAAG-CAATCGTAACATTACA 717
QY 2226 gcaagcagcatgcagcatc 2245
Db 718 G-AAGCAGCATGCAGAAATAC 736
RESULT 9
LOCUS BG217686 782 bp mRNA linear EST 21-APR-2001
DEFINITION RST37400 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG217686
VERSION BG217686.1 GI:13743707
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 782)
AUTHORS Harrington,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
,J., Dandzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scajin@atersys.com
High quality sequence stop: 473.
FEATURES
source
1. .782
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression'
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
BASE COUNT 243 a 130 c 144 g 245 t
ORIGIN
Query Match 17.8%; Score 637; DB 10; Length 782;
Best Local Similarity 95.5%; Pred. No. 3.3e-119;
Matches 677; Conservative 0; Mismatches 30; Indels 2; Gaps 2;
QY 2038 agtgttgaagatgtagtgatcttgaacacctccagacacagaacacacacacacac 2097
Db 2 AGTGTTTGAAAGATGTATGGTGATCTTGAAACCTCCAGACACAGAAACCTTCTAGCAAA 61
QY 2098 ttcagggaagttgtctacactcagctgagctattttcagcaacttgattgagcaaa 2157
Db 62 TTCAGGGGAAGTTGTCTACACTCAGCTGCGATTTTTCAGCAACCTTGATTGGCAAA 121
QY 2158 cggcctgtgccttatcttttgggtgagtgaaaaatttgagctagtgagcccaaatcgt 2217
Db 122 CGGCGCTGTGCCCTATCTTTGGTGGAGTGAATAAATTTGAGTTAGTGAAGCCAAATCGT 181
QY 2218 aacttacagaacgagcatgcagcatcaccttgctcttcttgattgcaaatagcattta 2277

Db	182	AACTTACAGAACGACGATGCACCATACCTGGCTCTTTGCTGA-TGCAAAATAGCGCTTTA	240
QY	2278	aaatgtgaatttggaatcagatgtctccattacttccagttaaagtggcatcatagtggt	2337
Db	241	AAATGTGAATTGGAAATCAGATGCTCCATTACTCCAGTTAAAGTGGCATCATAGGTGT	300
QY	2338	ttcctaagttttaagtccttggataaaataccaccagtgcttaccatctccaccatgaac	2397
Db	301	TTCTTAAGTTTAAAGCTTGGATAAAACTCCACCAGTGTCTACCATCTCCACCATGAAC	360
QY	2398	tctgttaagaagcttcatttttgtatattccgctcttttcttcttcattccctgtctt	2457
Db	361	TCGTGTTAAGAAGCTTCATTTTGTATATTCGGCTCTTTTCCTTCATTTCCCTGCTCT	420
QY	2458	ctgcataatcatgcctcttctgtaagtaattcaagcataagatctctgggaataataaaatc	2517
Db	421	CTGCATAATCATGCCCTCTCTTGCTTAAGTAATCAAGCATATGATCTTTGGAATAATAAATC	480
QY	2518	acaattctaggagaaagaataaaattgtttattttcccaagtctcttggccatgatgatac	2577
Db	481	ACAATCTTAGGAGAAAGATAAAATTCGCTATTTTCCAGTCTCTTGCCCATGATGATAC	540
QY	2578	ttatgtataaaacaaataataattttaaaccctgaagataaaattagaagaaattgtgc	2637
Db	541	TTATGATTTAAGACAAATTCATTTTAAACACCTGAATATAAATTAAGAANAATTGGGC	600
QY	2638	acctccacaaacatacaciaa-gtttaaagtttggtatcttttctcagcaggtatcagt	2696
Db	601	ACCTCCACAAACATACAAAGTCTAAAGTTGGGATCTTTTCCCTACCAGGTATAATG	660
QY	2697	tgtataatagttaattggggccaaatgcgaacgaaaaatgaagcagc	2745
Db	661	TGAATAATGAATTTGGGGCCAAATTCGAAACGAAATAATGAATCCGC	709
RESULT 10			
BG291485			
LOCUS			
602387373F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4516123 5',			
mRNA sequence.			
ACCESSION			
BG291485			
VERSION			
BG291485.1			
GT:13049409			
KEYWORDS			
EST.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1. (bases 1 to 936)			
NIH-MGC http://mgc.nci.nih.gov/.			
AUTHORS			
National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE			
Unpublished (1999)			
JOURNAL			
Contact: Robert Strausberg, Ph.D.			
COMMENT			
Email: cgapbs-r@mail.nih.gov			
Tissue procurement: ATCC			
cDNA library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLNL0407 row: a column: 20			
High quality sequence stop: 703.			
Location/Qualifiers			
1. .936			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:4516123"			
/clone_lib="NIH_MGC_93"			
/tissue_type="transitional cell papilloma, cell line"			
/lab_host="DH10B (phage-resistant)"			
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;			
Site_2: SalI; Cloned unidirectionally; Oligo-dr primed.			
Average insert size 1.7 Kb. Library enriched for			

FEATURES
source

full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."			
BASE COUNT 299 a 140 c 201 g 296 t			
ORIGIN			
Query Match 17.7%; Score 633.8; DB 10; Length 936;			
Best Local Similarity 93.6%; Pred. No. 1.5e-118;			
Matches 740; Conservative 0; Mismatches 37; Indels 14; Gaps 7;			
QY	2775	ctgtaattgaatttggtgcttcaatgattattttattttattgtacttttttcattattg	2834
Db	1	CTGTAAATTGAATATTGTGGCTTCATATGATATATTTATATTGACTTTTTTCATTATTG	60
QY	2835	atggtttgactttaaagagaaattccatagtttttaataatccacagaagtgaacaat	2894
Db	61	ATGGTTTGGACTTTAATAAGAGAAATTCATAGTTTTTAATATCCAGAAAGTGAGACAAT	120
QY	2895	ttgaacagtgtattcttagaataaacatacacactaaactgaacagaagtgaatgttatata	2954
Db	121	TTGAACAGTGTATTCTAGAAAAACAATACACTAACTGAACAGAAAGTGAATGCTTATATA	180
QY	2955	ttatgatagccttaaacctttttcccttaatgccttaactgtcaaatataataaaccttt	3014
Db	181	TTATGATAGCCTTAACCTTTTTCCTCTAATGCCTTAACTGTCAAATAATATTAACCTTT	240
QY	3015	taaagcataggactatagtcagcatgctagactgagaggttaaacactgatgcaattagaa	3074
Db	241	TAAAGCATAGGACTATAGTCAGCATGCTAGACTGAGAGGTAAACACTGATGCAATTAGAA	300
QY	3075	caggctactgctgctcagtggtttaacactatgttttagctgctgtttatgctataaaagt	3134
Db	301	CAGGTACTGATGCTGCTCAGTGTTTAAACACTATGTTTGTAGCTGTGTTATGCTATAAAAGT	360
QY	3135	caatattacacactagctagctgctcctcatgtaactcccaaaaaaacagagattica	3194
Db	361	CAATATTACACACTAGCTAGTCTGCTCATGTAACTCCAAAGAAACAGGATTTCA	420
QY	3195	ttaagtgcattgaattgggatatttctctaaagttaactcatattgtctcttctgtgaatgc	3254
Db	421	TTAAGTGCATTGAATGCTGGCTATTCTCTAAAGTTACTCATATTGCTTGTGTAATGC	480
QY	3255	aatgcctgcagatttatgagcgtctattttattttctctgtcattacttttaaaccttt	3314
Db	481	AATGCCGTGCAGATTATGTGGCTGCTATTATTTATTTCTGTGCATTACTTTTAACACCTT	540
QY	3315	aaagggaagc-aaacatttctcttcagctgactggcaatggccctttaactcaat	3373
Db	541	AAAGGGAAGAACAAACATTTCTCTTTCAGCTGACTGGCAATGGCCCTTTAACTGCAAT	600
QY	3374	aggaagaaaaaa-aaaaagggtttgtgaaaaattggtgataactggcacttaaga	3427
Db	601	AGGAAGAACACAAACCAACCAACGGGTTTGGTGGAAACTTGGTGATACTGGCACTTAAGA	660
QY	3428	tcgaaaagaaa-tttctgtatacttgatgccttaagaatgcccaaaagctgcccaagctct	3486
Db	661	TCGAAAAGAAATTTTCTGTATATCTTGTATGCTT-AGATGGCCCAAAGGTGCCAAGTCT--	717
QY	3487	gaagacttaagatagcgaataacttactacaatactactagttttttgtagagttta	3546
Db	718	-GAAGATTTAGGATAGGAGT-ATGCTTACTACTAC-ATACTACTGAGTTTGGGAAGTTAC	774
QY	3547	acatttgataa 3557	
Db	775	ATTGGTAATAA 785	
RESULT 11			
BI557167			
LOCUS			
603239140F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292152 5',			
DEFINITION			
mRNA sequence.			
ACCESSION			
BI557167			
LOCUS			
913 bp mRNA linear EST 05-SEP-2001			
BI557167			
LOCUS			
603239140F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292152 5',			
DEFINITION			
mRNA sequence.			
ACCESSION			
BI557167			

Db	421	TAAGTGAAGGACACAGAGATCTTAGCCCTGCAGAGGCAACAACATCACATGGGGACCCGGT	480
Qy	851	atattgaggtttcaaaagcaacaggtgaagattcccttaaaattgcttggttactcca	910
Db	481	ACATTGAGGTTTACAAGACCAACAGGCGAAGATTTCCCTTAAATTTCCGGAGGGACATCCA	540
Qy	911	atgaggtagccagttctctctccaaggaaaaatcaagtcattgttcgcattgccttactcg	970
Db	541	ATGAGGTAGCCCACTTCTCTCCAAGGAAAACCAAGTCATCGTGCATGAGGGGGCTCC	600
Qy	971	ctttcacggccacagctgaagaagtggtgcccctcttttgacacagcattgccttactcg	1030
Db	601	CTTTCACGGCCACGGCTGAAGAAGTGTGGCTCTTTTGGACACACTGCCTCCATCATG	660
Qy	1031	ggggaaggaagcgatcct-ctttgtcacctaccacagatggtagggccaa-caggggacgc	1088
Db	561	GAGGGAAGGAAGGATCTCTCTGTTACCTACCCAGATGGTAGGCCACACCGGAGATGC	720
Qy	1089	ttttgctctcttgcctgtgaggaatgcacagaatgcgttgaggaaagcataaagactt	1148
Db	721	TTTTTGTGCTCTTGTGTTGAGGAGTATGCTCAGACGCGCTGAGGAAGCACAAAGGAGTT	780
Qy	1149	gttgggttaaaagatacatt-gaacctttcaggagacacagcagctgaagttcagcaggtgc	1207
Db	781	GCTGGGTAAAAGATACCTTCGAGCTCTTCAGGACACAGGTGCCGAAAGTCAGCAAG-TG	838
Qy	1208	tgaatgatctctcctggccctcttcacttcacattccaacc	1249
Db	839	TTAACCGATTCTCTCTCAG-CCCTCTGAATCCACTTCCAACC	879
RESULT	12		
LOCUS	602382018F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:499644 5'	960 bp	linear EST 21-FEB-2001
DEFINITION	mRNA sequence.		
ACCESSION	BG287429		
VERSION	1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 960)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLAM10364 row: c column: 05 High quality sequence stop: 632.		
FEATURES	Location/Qualifiers		
source	1..960		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:499644"		
	/clone_lib="NIH_MGC_93"		
	/tissue_types="transitional cell papilloma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/notes="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
BASE COUNT	326 a 183 c 151 g 299 t 1 others		


```
Db 659 AAAACATACTAACTTTAAACCCCTTGAAGAAAAATTAGAGAAATTTTGGCCCCCTTC 718
Qy 2645 acaaacatacaaggtttaaaagtttg---atcttttctcagcagtgatcagttgtaa 2701
Db 719 CCAACATACAAAGGCTAAACGTTTGATCTCTTTTCTAAGCAGGTTCCCTGTTTA 778
Qy 2702 ----ataatgaattaggggccaaaatgcaaaacgaaaaaatg 2738
Db 779 CATTAATGAATATAGGGCCCAAAATAGAAAACCGATAAG 819

RESULT 14
BE870387 904 bp mRNA linear EST 20-OCT-2000
LOCUS 60144757f1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3851705 5',
DEFINITION mRNA sequence.
ACCESSION BE870387
VERSION BE870387.1 GI:10319163
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9572 row: m column: 18
High quality sequence stop: 622.
FEATURES
    source
        1..904
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3851705"
            /clone_lib="NIH_MGC_65"
            /tissue_type="adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
            Average insert size 1.8 kb. Library constructed by Life
            Technologies."
BASE COUNT 283 a 142 c 188 g 291 t
ORIGIN

Query Match 17.2%; Score 618; DB 10; Length 904;
Best Local Similarity 95.1%; Pred. No. 2.4e-115;
Matches 702; Conservative 0; Mismatches 30; Indels 6; Gaps 6;

Qy 2753 agttagtaatttcagtttgactgaattgaattatgtgcttcattgtattttta 2812
Db 1 AGTTAGTAATTTCTAGTTTGAACGTAAATGAATATTTGGCTTCATATGATTATTTA 60
Qy 2813 tatgtactttttcattatgatgttgactttaagaagaaatccatagttttt 2872
Db 61 TA-TGTACTTTTTTCATTATTGATGGTTGGACTTTAATAAGAGAAATCCATAGTTTTT 119
Qy 2873 aatatccagaagtgaacaatttgaacagtgattcttagaaaaacatacactaaactgaa 2932
Db 120 AATATCCCAAGAGTGAGACAAATTGAACAGTGATTCTTAGAAAACAAATACACTAACTGAA 179
Qy 2933 cagaagtgaatgcttatatatattatgatagccttaaaccttttccctaatgccttaa 2992
Db 180 CAGAAGTGAATGCTATATATATTATGATAGCCTTAAACCTTTTCCCTCTAATGCGCTAA 239
```

```
Qy 2993 ctgcaaaaataattataacctttttaagcataggactatagtcagcatgctagactgagag 3052
Db 240 CTGTCAAATANTATTAACCTTTTAAGCATAGGACTATAGTCAGCATGCTAGACTGAGAG 299
Qy 3053 gtaaacactgatgcaattagAACAGGtactgatgctgcagtggtttaacactatgttag 3112
Db 300 GTAAACACTGATGCAATTAGAACAGGTACTGATGCTGTCAG-GTTTAAACACTATGTTTAG 358
Qy 3113 ctgtgttatgctataaaagtcaaatattagacacactagctagctactgctcatgtaa 3172
Db 359 CTGTGTTTATGCTATAAAAGTGCAATATTAGACACTAGTAGTCTGCTCCCTCATGTAA 418
Qy 3173 ctcaaaagaaacagagatttcatttaagtgcattgaatggatatttctctaaagttaact 3232
Db 419 CTCCAAAGAAACAGGATTTCATTAAAGTGCA-TGAATGGGATATTCTCTAAGTTACTC 477
Qy 3233 atattgctctttgcttgaatgcaatgccgtgcagatttatgaggctgctattttattt 3292
Db 478 ATA-TGTCCTTTTGCTTGAATGCAATGCCGTGCAGATTTATGAGGCTGCTATCTTTATTTT 536
Qy 3293 ctgtgctattcttaaacaccttaaggagagaagcaaacatttccctcttca-gctgactg 3351
Db 537 CTGTGCATTACTTTAACACACCTTAAAGGGAGAGCAACACATATCCTTCTCAGGCTGACTG 596
Qy 3352 gcaatggcccttttaactgcaataggaagaaaaaaaggtttgtgtaaaattggtga 3411
Db 597 GCAATGGCCCTTTTAAGTCAATAGGACGAAACACACACAGGCTCTGTGTGAACATGGGTG 656
Qy 3412 taactggcacttaagatcgaaaagaatttctgtatactatgcttgaagatgcccaaa 3471
Db 657 ATAATGCACTTTTAAGATCGACAAG-AGTTTCTGTATTACTTGATGGCTTAAGATGCAAAAG 715
Qy 3472 gctgcccacagctctgaa 3489
Db 716 GTGGCCCAAGGTTCTGAA 733

RESULT 15
BG260588 890 bp mRNA linear EST 13-FEB-2001
LOCUS 602372144f1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4480163 5',
DEFINITION mRNA sequence.
ACCESSION BG260588
VERSION BG260588.1 GI:12770404
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10313 row: g column: 12
High quality sequence stop: 654.
FEATURES
    source
        1..890
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4480163"
            /clone_lib="NIH_MGC_93"
            /tissue_type="transitional cell papilloma, cell line"
            /lab_host="DH10B (phage-resistant)"
```

/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT		306 a	125 c	173 g	286 t
ORIGIN					
Query Match		17.2%	Score 615.8;	DB 10;	Length 890;
Best Local Similarity		97.58;	Pred. No. 6.7e-115;		
Matches 668;		Conservative 0;	Mismatches 12;	Indels 5;	Gaps 4;
QY	2571	tgatatcttattgattataaaacaaataattttaaacaacacctgaagataaattagaagaa	2630		
Db	1	TGATATCTTATGATTAAAAACAATTAATTTAAAAACACCTGAAGATATATTAGAAGAA	60		
QY	2631	atttgacacctccacaaacacatacaaaagttaaagtttggatcttttctcagcaggt	2690		
Db	61	A-TGTGCACCCCTCCACAAAACATACAAAGTTTAAAGTTTGGATCTTTTCTCAGCAGGT	119		
QY	2691	atcagttgtaataatgaattaggcgccaaatgcacaaacgaaaaatgaagcagctacat	2750		
Db	120	ATCAGTTGTAATTAATGAATTAGGGGCCAAATGCAAAACGAAAAATGAAGCAGCTACAT	179		
QY	2751	gtagttagtaattctctagtttgaactgtaattgaatattgtggcttcataatgtattatt	2810		
Db	180	GTAGTTAGTAATTTCTAGTTGAAGTGAATTTGAATTTGTGGCTTCATATGATTATTT	239		
QY	2811	tatatgtactttttcattatgatggtttggactttaagaagaaaattccatagttt	2870		
Db	240	TATATTGTACTTTTTCATTATTATGATGG-TTGGACTTTTAATAAGAGAAATTCCTAGTTT	298		
QY	2871	ttaatacccagagtgagacaatttgaacagtgatttctagaaaaacaatacacataactg	2930		
Db	299	TTAATATCCCAGAGTGAGACAATTTGAACAGTGATTTCTAGAAAAACAATACACTAACTG	358		
QY	2931	aacagaagtgatgcttatatatattatgatagccttaaaccttttccctcctaagtcctt	2990		
Db	359	AACAGAAGTGAATGCTTATATATATATATATGATAGCCTTAAACCTTTTCCCTAATGCGCTT	418		
QY	2991	aactgtcaataattataaaccttttaaagcataggaactatagtcagcatgctagactgag	3050		
Db	419	AACGTGCAATAATATAAOCCTTTTAAAGCATAGGACTATAGTCAGCATGCTAGACTGAG	478		
QY	3051	aggtaaacactgatgcaattagaaacaggtactgatgctgcagtggtttaacactatgttt	3110		
Db	479	AGGTAAACACTGATGCAATTAGAACAGGTACTGATGCTGTCAGTGTAAACACTATGTTT	538		
QY	3111	agctgtgtttatgctataaaagtgcaatatagacactagctagctgctcctcatgt	3170		
Db	539	AGCTGTGTTTATGCTATAAAAGTCAATATTAGACACTAGCTAGTACTGGTG-CTCATGT	597		
QY	3171	aactccaaagaaacacagatttccattaaagtcattggaattgaatttctcctaagttac	3230		
Db	598	ACTCCCAA--GAAACCGGATTTCAATTAAGTGCATTAAGTGAATTGCTCTAAGTTAC	655		
QY	3231	tcattatgctctttgcttgatgca	3255		
Db	656	TCATATTGGTCCCTGGTTGAATGCA	680		

Search completed: August 15, 2002, 08:56:29
Job time: 15240 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 06:04:49 ; Search time 383.11 Seconds
(without alignments)
2298.546 Million cell updates/sec

Title: US-09-697-206A-1

Perfect score: 3585

Sequence: 1 cttttggatcactgctgg.....gctgttaatctcaaaaaa 3585

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCITUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
c 1	305	8.5	351	4	US-09-328-111-380	Sequence 380, App
c 2	55.8	1.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
c 3	49.8	1.4	7218	1	US-08-232-463-14	Sequence 14, Appl
c 4	47.8	1.3	602	1	US-08-764-100-8	Sequence 8, Appl
c 5	47.8	1.3	643	1	US-08-764-100-7	Sequence 7, Appl
c 6	47.8	1.3	2993	1	US-08-764-100-2	Sequence 2, Appl
c 7	47.8	1.3	2993	1	US-08-764-100-2	Sequence 10, Appl
c 8	47.8	1.3	3001	1	US-08-764-100-1	Sequence 1, Appl
c 9	44.6	1.2	837	4	US-08-998-416-288	Sequence 288, App
c 10	43.8	1.2	636	4	US-08-998-416-1137	Sequence 1137, App
c 11	43.8	1.2	642	1	US-08-764-100-13	Sequence 13, Appl
c 12	43.8	1.2	3000	1	US-08-764-100-9	Sequence 9, Appl
c 13	43.8	1.2	5852	1	US-07-867-106-2	Sequence 2, Appl
c 14	43.8	1.2	7218	1	US-08-232-463-14	Sequence 14, Appl
c 15	41.6	1.2	615	4	US-08-998-416-186	Sequence 186, App
c 16	41.6	1.2	1431	4	US-09-316-083-2	Sequence 2, Appl
c 17	41.4	1.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
c 18	40.8	1.1	10968	2	US-08-680-327-2	Sequence 2, Appl
c 19	40.8	1.1	10968	4	US-09-228-246-1	Sequence 1, Appl
c 20	40.2	1.1	1939	1	US-07-715-751B-2	Sequence 2, Appl
c 21	40.2	1.1	5506	4	US-09-004-838-93	Sequence 93, Appl
c 22	40	1.1	51952	3	US-08-947-823-1	Sequence 1, Appl
c 23	39.8	1.1	3095	6	5231168-1	Patent No. 5231168
c 24	39.8	1.1	5852	1	US-07-867-106-2	Sequence 2, Appl
c 25	39.6	1.1	636	4	US-08-998-416-1137	Sequence 1137, App
c 26	39.6	1.1	731	1	US-08-451-405A-2	Sequence 2, Appl
c 27	39.6	1.1	827	4	US-08-998-416-535	Sequence 535, App

```

c 28 39.4 1.1 665 4 US-08-998-416-937 Sequence 937, App
c 29 39.2 1.1 6243 2 US-09-056-075-1 Sequence 1, Appli
c 30 39 1.1 1236 6 5202336-4 Patent No. 5202336
c 31 39 1.1 5630 2 US-08-937-931-1 Sequence 1, Appli
c 32 39 1.1 5630 4 US-09-285-502-1 Sequence 1, Appli
c 33 39 1.1 5630 4 US-09-709-126-1 Sequence 1, Appli
c 34 38.6 1.1 615 4 US-08-998-416-186 Sequence 3, Appli
c 35 38.6 1.1 6124 4 US-08-213-419B-3 Sequence 3, Appli
c 36 38.6 1.1 6243 2 US-09-056-075-1 Sequence 1, Appli
c 37 38.2 1.1 15766 4 US-09-338-907-73 Sequence 73, Appli
c 38 38.2 1.1 15766 4 US-09-218-207-73 Sequence 73, Appli
c 39 38.2 1.1 37950 4 US-09-338-907-183 Sequence 183, App
c 40 38.2 1.1 37950 4 US-09-218-207-183 Sequence 183, App
c 41 38 1.1 767 4 US-08-998-416-472 Sequence 472, App
c 42 38 1.1 782 4 US-08-998-416-224 Sequence 224, App
c 43 37.8 1.1 662 4 US-08-998-416-185 Sequence 185, App
c 44 37.8 1.1 701 4 US-08-998-416-701 Sequence 701, App
c 45 37.8 1.1 724 4 US-08-998-416-683 Sequence 683, App

```

ALIGNMENTS

```

RESULT 1
US-09-328-111-380/c
; Sequence 380, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 380
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-328-111-380

```

Query Match 8.5%; Score 305; DB 4; Length 351;
Best Local Similarity 96.3%; Pred. No. 5, 2e-71;
Matches 338; Conservative 0; Mismatches 0; Indels 13; Gaps 2;

```

Qy 1388 gtactcattgggtccacatggtttgaatcacccaggccgcacagatgagatgccttta 1447
Db 350 GTACTCATGGGTTTCACATGGTTTGAATCACAGGGCCGCCATCAGGAGATGCTTTA 291
Qy 1448 tccagatgaagtctcggacagacatttatgctgcacagaagtgtcataaaaaaaca 1507
Db 290 TCCAGATGAAGTCTCGGACAGACATTTATGCGTGCACAGAAGTGTTCAT-AAAAAACA 232
Qy 1508 tgaaggacagatatgttaagtcttcagtggttcagctgagagatgaactttgtgttaa 1567
Db 231 TGAGGACAGACATATGTTGAAGTCTTTCAGTGTTCAGTGCAGGAGATGAACTTTGTGTTAA 172
Qy 1568 tggggggcactttaaatgaaatggcttattcccccccg-----ccatgcctgt 1615

```

Qy	2533	agaataaaatgttattttcccgagtcctcttgcccgatgatatactcttcatgattaaaaaca	2592
Db	6442	AAATATATTTTTTCCTGTTATTTATTTATCATATTTTTTTTTTGATGCTATATATATTATTATA	6501
Qy	2593	aattaaatttttaaacacacctggaagataaaattagaagaattgtgcac-cctcccaaaac	2651
Db	6502	TAATAAATTATAATAATATACACAAAAAATTAAATAAATAATATACATACATATAATAATA	6561
Qy	2652	atacaaggtttaaaagtttggatcttttctcagcagggtatccagttgttaaataatgaatt	2711
Db	6562	ATACACAATACAAAGAATATGTACTATATCAATTATATATATGAATATATAAATAT	6621
Qy	2712	agggggccaaatgcgaaacgaaataaagaagcagctacatgtagttagtaattctctagttt	2771
Db	6622	GATAGATAATATAGATAGAGAGAAACGAAGAACAATATTGTCTCTCTTTGTTATC---TCT	6678
Qy	2772	gaactgtaattgaatttggcttccatgtattattttatttatttattgtaacttttttcatta	2831
Db	6679	AAATATATATATATATAATAAATTAAAATAAAGTCAAAAAAATATACATATATTAATG	6738
Qy	2832	ttgatggttggaactttaagaagagaaaattccatagtttttaaatatccagaagtgcagac	2891
Db	6739	TTAAATAAATAATATAAACAACAGCTTGCATATATACTTTTTTATATGTTGTATTTTCGT	6798
Qy	2892	aatttgaacagtgattcttagaaaaaatcacactaacctgaacagaagtgaatgccttat	2951
Db	6799	ATTTTTTTTTTCTCAATTTATAATTTTACTTTAATAAATAAAACATAAAAAAATAATATAT	6858
Qy	2952	atattatgatgccttaaa	2970
Db	6859	ATAAATAAATAAGATAAA	6877

```

RESULT      3
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:

```



```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
; US-08-232-463-14

Query Match 1.4%; Score 49.8; DB 1; Length 7218;
Best Local Similarity 6.3%; Pred. No. 0.0033;
Matches 21; Conservative 180; Mismatches 132; Indels 0; Gaps 0;

QY 1593 ctatccaccacccatgctgctcctccctccctacacatttccagctcctgctcagt 1652
Db 1142 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1201

QY 1653 tattctacagaagctgccatttaccagccctctgtgatttgaatccacgagcactgca 1712
Db 1202 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1261

QY 1713 gcctccacagctactaccacgacgacgcacgcctcttcatgaactacacagcgtacta 1772
Db 1262 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1321

QY 1773 tccagcccccaggctgcctaatagcttgctgctactccctacagctgctacttag 1832
Db 1322 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1381

QY 1833 cgtgtccctccacagctggcagcgtggtcagaatgcaggcctgcctacaactg 1892
Db 1382 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1441

QY 1893 agttaaggaaattcttaactcttccaaggtta 1925
Db 1442 ATTCTTCTATCTCTTAACTACTTGTCATAGATA 1474

RESULT 4
US-08-764-100-8
; Sequence 8, Application US/08764100
; Patent No. 5773700
; GENERAL INFORMATION:
; APPLICANT: van Grinsven J., Martinus Q.
; APPLICANT: De Haan, Petrus T.
; APPLICANT: Gielen L., Johannes J.
; APPLICANT: Peters, Dirk
; APPLICANT: Goldbach, Robert W.
; TITLE OF INVENTION: Improvements in or Relating to Organic
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Agro, Inc
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,100
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,064
; FILING DATE:
; APPLICATION NUMBER: US 08/032,235
; FILING DATE: 17-MAR-1993
```

```
;
; APPLICATION NUMBER: GB 9206016.9
; FILING DATE: 19-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 57737001s, Allen E.
; REGISTRATION NUMBER: 34,490
; REFERENCE/DOCKET NUMBER: 137-1061
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 354-3592
; TELEFAX: (415) 857-1125
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-764-100-8

Query Match 1.3%; Score 47.8; DB 1; Length 602;
Best Local Similarity 47.8%; Pred. No. 0.0035;
Matches 139; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 2570 atgatctcttatgattaaacaaattaaatttttaaaccacctgaagataaattagaaga 2629
Db 85 ATCATATGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 144

QY 2630 aattgtcacccctccacaaacatacaaaagtttaaaagtttgatctttttctcagcagg 2689
Db 145 TAAATAAAACCAAAAGGATGCGCTTCGGGCACAAATTTGGTGTCTTTAATAATGCTTT 204

QY 2690 tatcagttgtaataatgaattagggggcccaaaatgcacaaacgaaatgaagcagctaca 2749
Db 205 AAAATGAATCTATTAGTAATATTAATACTTTAAATCCAATCTACTCACAATGGCCAAA 264

QY 2750 tgtagtagtaattcttagtttgaactgtaattgaaatattgtggtcttcattattatt 2809
Db 265 AATTGTATTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTT 324

QY 2810 ttattgttactttttctattgattggttggttgaccttaataagagaat 2860
Db 325 TTTTATTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTAT 375

RESULT 5
US-08-764-100-7
; Sequence 7, Application US/08764100
; Patent No. 5773700
; GENERAL INFORMATION:
; APPLICANT: van Grinsven J., Martinus Q.
; APPLICANT: De Haan, Petrus T.
; APPLICANT: Gielen L., Johannes J.
; APPLICANT: Peters, Dirk
; APPLICANT: Goldbach, Robert W.
; TITLE OF INVENTION: Improvements in or Relating to Organic
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Agro, Inc
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,100
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
```



```
; NAME: No. 5773700ris, Allen E.
; REGISTRATION NUMBER: 34,490
; REFERENCE/DOCKET NUMBER: 137-1061
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 354-3592
; TELEFAX: (415) 857-1125
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-764-100-9

Query Match          1.2%; Score 43.8; DB 1; Length 3000;
Best Local Similarity 49.1%; Pred. No. 0.083;
Matches 143; Conservative 0; Mismatches 147; Indels 1; Gaps 1;

QY 2570 atgatatcttatgattataaaacacaaatttaaaacacacgtgaagataaattagaaga 2629
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1493 ATCATATGAATTAATAATAACATAATAATAATAATAATAATAATAATAATAATAATA 1434

QY 2630 aatttgaccctccacaaacatacaaaagttaaagtttaaaagtttgatctttttctcagcagg 2689
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1433 TAAATTAATAAACCAAAAGGATGGCTTCGGGCACAA-TTTGGTTGCTTTTAATAATGCTT 1375

QY 2690 tatcagttgtaaatgaattagggccaaatgcacaaacgacaaatgaagcagctaca 2749
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1374 TAAAGATGATATTAGTAAATTAATAATTAATTAATTAATTAATTAATTAATTAAT 1315

QY 2750 tgaagttagtaattctctagttgaaactgtaattgaaactgtaattggttcacatgtattatt 2809
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1314 AATTTGTAATGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1255

QY 2810 ttatatgtactttttcattatgattggtttgggactttaataagagagaat 2860
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1254 TTTTATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1204

RESULT 13
US-07-867-106-2
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
```

```
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
;
US-07-867-106-2

Query Match          1.2%; Score 43.8; DB 1; Length 5852;
Best Local Similarity 44.8%; Pred. No. 0.11;
Matches 168; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 2583 attaaaaacaaatttaaaacacacgtgaagataaaattagaagaaattgtgcacct 2642
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5387 ATTTATAAAATTGAAGTTCATCAAGATATATAGATAAATTATTTAATTATTTGAATTTT 5446

QY 2643 ccacaaacacatacaaaagtttaaaagtttggatctttttctcagcaggtatcagttgtaaa 2702
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5447 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 5506

QY 2703 taatgaattaggggccaaatgcacaaacgacaaatgaagcagctacatgtagttagtaaat 2762
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5507 TCTCGTCATGATTTTAAATAATAAAATCGATACATAAATTTTAAAAAACCCTTTACATT 5566

QY 2763 ttctagtttgaactgtaattgaaatattgtggcttcacatgtattattttatattgtactt 2822
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5567 TTTTATTTTAAATCCAAATTTATACATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5626

QY 2823 ttttcattatgattggtttggactttaataagagaaattccatagtttttaataatcccaag 2882
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5627 TTAATTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5686

QY 2883 aagtgagacaattgacagtgatcttclagaaaaacacacactaacctgaacagagagtgaa 2942
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5687 ATAAATAAAAAATACAAATATAAGTAAAAAACAACAAACAAATACATATATATAAAAAA 5746

QY 2943 tgcctatatattat 2957
   ||||| ||||| |||||
Db 5747 TACAAATAACAATA 5761

RESULT 14
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



```
Db 178 ATAAATATTAAATGAATGAATTAAGTAAATATATATAATTAATAAAGTATTAAT 119
QY 2893 atttgacagtgatttctagaaaaacaatacactaacctgaacagaagtgaatgcttatata 2952
Db 118 AATCAAAATTAATTAATTAATAATGATAAATAAGTTTAAATAAATTAATAACTTAAA 59
QY 2953 tattatgatagccttaaacctttttccctcctaatgccttaactgtcaataaattataac 3010
Db 58 TATTATAATAAAAAAGTTTATATTAATCTTTATAAATTAATTAATATATATATTAGATC 1
```

Search completed: August 15, 2002, 14:12:13
Job time: 29244 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2002, 06:20:49 ; Search time 2423.18 seconds
(without alignments)
2540.108 Million cell updates/sec

Title: US-09-697-206A-1
Perfect score: 3585
Sequence: 1 cttttggatcactgctgg.....gcctgttaatctcaaaaaa 3585

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-embl/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-embl/NA2001C.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3585	100.0	3598	22 AAF85701	Human cancer relat
2	2521.4	70.3	2882	21 AAA12412	cDNA encoding a hu
3	2360	65.8	2731	22 AAF98702	Human ovarian-canc
4	2316.8	64.6	2468	22 AAH33714	Human colon cancer
5	1986.2	55.4	2261	21 AAC77930	Human cancer assoc
6	1324.8	37.0	1624	22 AAH14740	Human cDNA sequenc
7	587.4	16.4	1585	21 AAC59113	Human secreted pro
8	555.6	15.5	740	20 AAZ16154	Human gene express
9	524	14.6	745	20 AAZ16155	Human gene express

c	10	505.4	14.1	565	22	AAH12166	Human cDNA clone (
	11	498.4	13.9	503	22	AAF68133	Human lung tumour
	12	498.4	13.9	503	22	AAF68134	Human lung tumour
	13	441.4	12.3	700	21	AAA01641	Human colon cancer
	14	392.4	10.9	745	22	AAH07171	Human cDNA clone (
	15	351	9.8	353	22	AAI29269	Colon tumour relat
	16	343.8	9.6	396	22	AAF94890	Human ovarian canc
	17	322.8	9.0	329	22	AAI29175	Colon tumour relat
c	18	305	8.5	351	21	AAZ80296	Human colon cancer
	19	272.4	7.6	411	22	AAZ80296	cDNA encoding nove
	20	272.4	7.6	411	22	AAZ80296	Human digestive sy
c	21	271	7.6	271	22	AAZ80296	Human digestive sy
c	22	271	7.6	271	22	AAZ80296	Human breast canc
	23	231	6.4	300	20	AAZ14741	Human breast canc
	24	227	6.3	2115	23	ABL12241	Human gene express
c	25	223	6.2	294	22	AAF98661	Drosophila melanog
	26	166	4.6	241	22	AAZ46997	Human ovarian canc
	27	166	4.6	241	22	AAZ46997	Human breast canc
	28	156.6	4.4	260	21	AAA00145	Human breast canc
	29	153	4.3	153	22	AAZ39650	Human colon cancer
	30	153	4.3	153	22	AAZ39650	Genomic sequence #
	31	151.4	4.2	157	16	AAZ89054	Human digestive sy
c	32	129	3.6	155	21	AAA45873	Human gene signatu
	33	127.4	3.6	9102	23	ABL12240	Human secreted exp
	34	120.2	3.4	317	20	AAZ40462	Drosophila melanog
	35	118.2	3.3	339	23	AAZ58815	Human secreted pro
	36	116.8	3.3	427	21	AAZ58815	cDNA #1491 encodin
	37	79.8	2.2	242	22	AAZ15980	Human OREX ORF1560
	38	79.4	2.2	257	22	AAZ15980	Human breast canc
	39	66.4	1.9	2670	22	AAH14219	Human breast canc
	40	61.4	1.7	826	22	AAI97495	Human cDNA sequenc
	41	61.2	1.7	1036	21	AAZ00757	Human neuroblastom
	42	60.2	1.7	1034	21	AAZ10857	Human secreted pro
	43	60.2	1.7	2201	22	AAZ83093	Human secreted pro
	44	60	1.7	815	22	AAI96163	hNRP HI ovarian tu
	45	58.6	1.6	2234	23	AAZ66261	Human neuroblastom
							DNA encoding novel

ALIGNMENTS

RESULT	1
AAF85701	
ID	AAF85701 standard; cDNA; 3598 BP.
XX	
AC	AAF85701;
XX	
DT	10-DEC-2001 (first entry)
XX	
DE	Human cancer related protein 20P2H8 coding sequence.
XX	
KW	Human; cancer related protein 20P2H8; vaccine; chromosome 15q32-23;
KW	prostate cancer; bladder cancer; colon cancer; pancreatic cancer; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..2148
FT	/*tag= a
FT	/product= "20P2H8 alternative version"
FT	451..2004
FT	/*tag= b
FT	/product= "20P2H8"
XX	
PN	WO200131012-A1.
XX	
PD	03-MAY-2001.
XX	
PF	26-OCT-2000; 2000WO-US29477.
XX	
PR	28-OCT-1999; 99US-0162364.
XX	
PA	(UROC-) UROGENESYS INC.

XX Afar DEH, Raitano AB, Hubert RS, Mitchell SC, Jakobovits A;
XX WPI; 2001-308645/32.
DR P-PSDB; AAB81201, AAB60948.
XX 202H8 polynucleotides and polypeptides useful for diagnosing and
PT treating cancer, and for screening for screening for modulating
PT compounds. -
XX
XX
PS Claim 1; Fig 1; 111pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC human cancer related protein 20P2H8. The gene, which is found at
CC chromosome 15q32-23, is upregulated in cancers such as that of the
CC prostate, bladder, colon and pancreas. The sequences can be used to
CC diagnose and treat these cancers, and to vaccinate against them. The
CC present sequence is the coding sequence of the invention.
XX
SQ Sequence 3598 BP; 1072 A; 741 C; 763 G; 1022 T; 0 other;

Query Match 100.0%; Score 3585; DB 22; Length 3598;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctttttggatcactgctggggccaccgggcccaggctaggctgagtgagaaggattg 60
Db 1 ctttttggatcactgctggggccaccgggcccaggctaggctgagtgagaaggattg 60

Qy 61 atctctgctgttctgaaagtctgagatctggccaaagaagggtgggacagttgcacgaa 120
Db 61 atctctgctgttctgaaagtctgagatctggccaaagaagggtgggacagttgcacgaa 120

Qy 121 gtgctagtagcagcagatcagttgaaactgacgagagactcaagaagaactaaata 180
Db 121 gtgctagtagcagcagatcagttgaaactgacgagagactcaagaagaactaaata 180

Qy 181 gacgtcgaagcctgctctcgcgcgcagctggaccagccctccgacagtttaaccag 240
Db 181 gacgtcgaagcctgctctcgcgcgcagctggaccagccctccgacagtttaaccag 240

Qy 241 taagtgaacatgaactgaatattgagtagggacttctctctgctgactgaggg 300
Db 241 taagtgaacatgaactgaatattgagtagggacttctctctgctgactgaggg 300

Qy 301 cagcttcagtcagcgaactcctcatcctgagcttccaaagaagaatgtactattact 360
Db 301 cagcttcagtcagcgaactcctcatcctgagcttccaaagaagaatgtactattact 360

Qy 361 gaatgcttctattctctttttgatcttcgaaagaattcaagaatgttgcctggttca 420
Db 361 gaatgcttctattctctttttgatcttcgaaagaattcaagaatgttgcctggttca 420

Qy 421 cctgatatgcaaaactggaactgtccaaatgacagagatttaaaattttgagaagagt 480
Db 421 cctgatatgcaaaactggaactgtccaaatgacagagatttaaaattttgagaagagt 480

Qy 481 agttcagctctcgataggagcctctcaagtgtgaagatagggaataatatttagca 540
Db 481 agttcagctctcgataggagcctctcaagtgtgaagatagggaataatatttagca 540

Qy 541 atgatttcagagccttaataacacagggttttcagatccagagagagtgaattacaagttt 600
Db 541 atgatttcagagccttaataacacagggttttcagatccagagagagtgaattacaagttt 600

Qy 601 gaaagtgaaacttcagcaagaatggaacttattgatataacacagtagtcagggcacga 660
Db 601 gaaagtgaaacttcagcaagaatggaacttattgatataacacagtagtcagggcacga 660

Qy 661 ggtttaccatggcagctcttcagatcaagatatgtcaagattcttcaaggactcaatt 720
Db 661 ggtttaccatggcagctcttcagatcaagatatgtcaagattcttcaaggactcaatt 720

Qy 721 gccaaaggaggtgcagcactttgtctgaatgctcaggggtcgaaggaaacggagaactctg 780
Db 721 gccaaaggaggtgcagcactttgtctgaatgctcaggggtcgaaggaaacggagaactctg 780

Qy 781 gttaggtttgtaagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 840
Db 781 gttaggtttgtaagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 840

Qy 841 gggaccgggtattgaggtttacaaagcaacagggtgagatttctttaaattctggt 900
Db 841 gggaccgggtattgaggtttacaaagcaacagggtgagatttctttaaattctggt 900

Qy 901 ggtactccaatgagtagccagtttcttccaaggaaatcaagtcattgttcgcatg 960
Db 901 ggtactccaatgagtagccagtttcttccaaggaaatcaagtcattgttcgcatg 960

Qy 961 cgggggtctcttccacggccacagctgaaagtggtggccttcttggacagcattgc 1020
Db 961 cgggggtctcttccacggccacagctgaaagtggtggccttcttggacagcattgc 1020

Qy 1021 cctattactgggaaaggaaggcactctcttgcacctaccagatggttaggcaca 1080
Db 1021 cctattactgggaaaggaaggcactctcttgcacctaccagatggttaggcaca 1080

Qy 1081 ggggacgcttttctctcttgcctgtgaggaatgacagaaatgcgtttgaggaagcat 1140
Db 1081 ggggacgcttttctctcttgcctgtgaggaatgacagaaatgcgtttgaggaagcat 1140

Qy 1141 aagaactgttgggtaaaagatacattgaactcttcagaggacacagcagcagcagcagcag 1200
Db 1141 aagaactgttgggtaaaagatacattgaactcttcagaggacacagcagcagcagcagcag 1200

Qy 1201 caggtgctgaatcagatctctcgcgcctctcattccacttccaaacccctccattatt 1260
Db 1201 caggtgctgaatcagatctctcgcgcctctcattccacttccaaacccctccattatt 1260

Qy 1261 ccagtactacctcagcaatttgcctccactacaaatgttagagactgtatagcgtctcga 1320
Db 1261 ccagtactacctcagcaatttgcctccactacaaatgttagagactgtatagcgtctcga 1320

Qy 1321 ggtttcctctatgcagccacaaatgcaggaactcctgattcctggggaggttcgcaaca 1380
Db 1321 ggtttcctctatgcagccacaaatgcaggaactcctgattcctggggaggttcgcaaca 1380

Qy 1381 gatattcgtactcagtggttcacatggtttgaatcaccagggcccccacagcagat 1440
Db 1381 gatattcgtactcagtggtttcacaatggtttgaatcaccagggcccccacagcagat 1440

Qy 1441 gcctttatccagatgaagctcgcggacagagcatttatggctgcacagaagtgtcataaa 1500
Db 1441 gcctttatccagatgaagctcgcggacagagcatttatggctgcacagaagtgtcataaa 1500

Qy 1501 aaaaacatgaaggacagatattgtgaagtcttcagtggttcagctgagagatgaacttt 1560
Db 1501 aaaaacatgaaggacagatattgtgaagtcttcagtggttcagctgagagatgaacttt 1560

Qy 1561 gtgttaattgggggcacttttaaatcgaaatggcttatccccaccgcctgctctcct 1620
Db 1561 gtgttaattgggggcacttttaaatcgaaatggcttatccccaccgcctgctctcct 1620

Qy 1621 cctctctacacatttccagctcctgctgagttatttctacagaagtgcatttaccag 1680
Db 1621 cctctctacacatttccagctcctgctgagttatttctacagaagtgcatttaccag 1680

Qy 1681 cctctgtagatttgaatccacagcagcagcagcagcagcagcagcagcagcagcagcag 1740
Db 1681 cctctgtagatttgaatccacagcagcagcagcagcagcagcagcagcagcagcagcag 1740

Qy 1741 actcagctcttcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1800
Db 1741 actcagctcttcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1800

QY 1801 cttgctactctccacagctgcttaattagcgtgtccctccacagcctggcagcgtg 1860
Db 1801 cttgctactctccacagctgcttaattagcgtgtccctccacagcctggcagcgtg 1860
QY 1861 gtcagaatgcaggcctggcctacataactgagtgtaaggaaattcttaactttctcaa 1920
Db 1861 gtcagaatgcaggcctggcctacataactgagtgtaaggaaattcttaactttctcaa 1920
QY 1921 ggttaccagtagcaccgagtaggacttatcacacaaaatgaccaggccaggactcta 1980
Db 1921 ggttaccagtagcaccgagtaggacttatcacacaaaatgaccaggccaggactcta 1980
QY 1981 cccaaagaatggtttgtatttaaggcccgagcagtgtagaacatccctcagaaaagaagt 2040
Db 1981 cccaaagaatggtttgtatttaaggcccgagcagtgtagaacatccctcagaaaagaagt 2040
QY 2041 gtttgaagatgtatggtagcttgaacacctccagacacagaaaactcttagcaaatc 2100
Db 2041 gtttgaagatgtatggtagcttgaacacctccagacacagaaaactcttagcaaatc 2100
QY 2101 aggggaagtgtgtcacactcagcgtcagtgattttcagcaaaacttgatggacaaacgg 2160
Db 2101 aggggaagtgtgtcacactcagcgtcagtgattttcagcaaaacttgatggacaaacgg 2160
QY 2161 gctgtgcttatcttttggtagtgtaaaatgttagctagtgaaagccaaaatcgtaac 2220
Db 2161 gctgtgcttatcttttggtagtgtaaaatgttagctagtgaaagccaaaatcgtaac 2220
QY 2221 ttacagcaagcagatgcagcacaactggctcttcttgctgattgcaaataggcatttaaa 2280
Db 2221 ttacagcaagcagatgcagcacaactggctcttcttgctgattgcaaataggcatttaaa 2280
QY 2281 tgtgaatttgaatcagatgtctccattacttccagttcaagtgagcattagtggtttc 2340
Db 2281 tgtgaatttgaatcagatgtctccattacttccagttcaagtgagcattagtggtttc 2340
QY 2341 ctaagttttaagttgtgataaaactccacagtgcttacatctccacatgaactct 2400
Db 2341 ctaagttttaagttgtgataaaactccacagtgcttacatctccacatgaactct 2400
QY 2401 gtttaggaagcttcaattttgtatattccgctcttcttcttcttcaattccctgtctctg 2460
Db 2401 gtttaggaagcttcaattttgtatattccgctcttcttcttcttcaattccctgtctctg 2460
QY 2461 cataatcatgctcttctgtcgaagtaattcaagcataaagatcttgaataataaaatcaca 2520
Db 2461 cataatcatgctcttctgtcgaagtaattcaagcataaagatcttgaataataaaatcaca 2520
QY 2521 atcttaggagaagaataaaattgtttatttccagtgctcttggccatgatgatatttta 2580
Db 2521 atcttaggagaagaataaaattgtttatttccagtgctcttggccatgatgatatttta 2580
QY 2581 tgattaaaaaacaataatttttaaaacacctgaagataaaattagaagaataatttgcacc 2640
Db 2581 tgattaaaaaacaataatttttaaaacacctgaagataaaattagaagaataatttgcacc 2640
QY 2641 cttcccaaaaacatacaagtttaaaagtgttgatcttttctcagcaggtatcagttgta 2700
Db 2641 cttcccaaaaacatacaagtttttaaaagtgttgatcttttctcagcaggtatcagttgta 2700
QY 2701 aataatgaattaggggcccataatgcataaagcgaataatgaagcagctacatagtagtagta 2760
Db 2701 aataatgaattaggggcccataatgcataaagcgaataatgaagcagctacatagtagtagta 2760
QY 2761 atttctagttgaactgttaattgaatttggcttccatgtatgtatttttatattgtac 2820
Db 2761 atttctagttgaactgttaattgaatttggcttccatgtatgtatttttatattgtac 2820
QY 2821 ttttttctattatgtatggttggacttttaataagagaataattccatagtttttaatatccc 2880
Db 2821 ttttttctattatgtatggttggacttttaataagagaataattccatagtttttaatatccc 2880
QY 2881 agaagtgcagaatttgaacagtgatttctagaaaaacaatacactaaactgaacagaagt 2940

Db 2881 agaagtgcagaatttgaacagtgatttctagaaaaacaatacactaaactgaacagaagt 2940
QY 2941 aatgcttatataattatgatagccttaaacctttttctcttaactgaacttaactgcaaa 3000
Db 2941 aatgcttatataattatgatagccttaaacctttttctcttaactgaactgcaaa 3000
QY 3001 aaattataaacttttaagcattagactatagtcagcagctagactgagaggttaaacac 3060
Db 3001 aaattataaacttttaagcattagactatagtcagcagctagactgagaggttaaacac 3060
QY 3061 tgaatcaattagaacaggtactgactgctgctgactgttttaacacactatgttttagctgtgtt 3120
Db 3061 tgaatcaattagaacaggtactgactgctgctgactgttttaacacactatgttttagctgtgtt 3120
QY 3121 atgctataaaagtgaataattatagacactagctagctagctgctcctcactgtaactccaag 3180
Db 3121 atgctataaaagtgaataattatagacactagctagctagctgctcctcactgtaactccaag 3180
QY 3181 aaaaacagatttccattgaatgctgatttgatatttcttaagttactcattatgtc 3240
Db 3181 aaaaacagatttccattgaatgctgatttgatatttcttaagttactcattatgtc 3240
QY 3241 ctttcttgtaatgcaatccgtgcagatttatgagctgctattttttttctgtgcat 3300
Db 3241 ctttcttgtaatgcaatccgtgcagatttatgagctgctattttttttctgtgcat 3300
QY 3301 tactttaacaccttaaaagggaagcaaacatttctcttcagctgactggcaatggcc 3360
Db 3301 tactttaacaccttaaaagggaagcaaacatttctcttcagctgactggcaatggcc 3360
QY 3361 ctttaactgcgaataggaagaaaaaaaagggttggtaaaattggtaactggca 3420
Db 3361 ctttaactgcgaataggaagaaaaaaaagggttggtaaaattggtaactggca 3420
QY 3421 ctttaagatcgaaaaaatttctgtatctactgctgcttaagatgcccaagctgccccaa 3480
Db 3421 ctttaagatcgaaaaaatttctgtatctactgctgcttaagatgcccaagctgccccaa 3480
QY 3481 agctctgaagaactttaagatagcaggtaagttaacttaacatactactgagttttgta 3540
Db 3481 agctctgaagaactttaagatagcaggtaagttaacttaacatactactgagttttgta 3540
QY 3541 gaggttacaactttgataataaaacttgcctgtttaatctcaaaaaa 3585
Db 3541 gaggttacaactttgataataaaacttgcctgtttaatctcaaaaaa 3585

RESULT 2

AAAL2412

ID AAAL2412 standard; cDNA; 2882 BP.

XX AAAL2412;

XX 25-JUL-2000 (first entry)

XX cDNA encoding a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
immune response; reproductive disorder; actinic keratosis;
KW arteriosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
KW trauma; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 293..2338

FT /*tag= a

FT /product= "RNA-associated protein"

XX PN W0200015799-A2.

XX FD 23-MAR-2000.
 XX PF 17-SEP-1999; 99WO-US21688.
 XX PR 17-SEP-1998; 98US-0156039.
 XX PR 22-SEP-1998; 98US-0158720.
 XX PR 04-NOV-1998; 98US-0186815.
 XX PR 08-APR-1999; 99US-0128660.
 XX PA (INCY-) INCYTE PHARM INC.
 XX XX
 PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
 XX WPI; 2000-271437/23.
 DR P-PSDB; RAY84440.
 XX XX
 PT New polypeptides and polynucleotides, useful for preventing and
 PT treating a disorder associated with increased or decreased expression
 PT of RNA associated proteins -
 XX
 PS Claim 9; Page 121-122; 131pp; English.
 XX
 CC The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used
 CC for diagnosis, treatment or prevention of cell proliferative,
 CC immune/inflammatory disorders, and reproductive disorders. Diseases
 CC and disorders which may be treated include actinic keratosis, skin
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis,
 CC mixed connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma.
 XX
 SQ Sequence 2882 BP; 755 A; 691 C; 671 G; 765 T; 0 other;

Query Match 70.3%; Score 2521.4; DB 21; Length 2882;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 2544; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

Qy 1 ctttttggatcactgtggtggccacgggcccagctagctgctgagtgagagagttg 60
 Db 323 ctttttggatcactgtggtggccacgggcccagctagctgctgagtgagagttg 382
 Qy 61 atctgtcttctgaaagtctggtgactggtcccaagaaggtggacagttgcacgaa 120
 Db 383 atctgtcttctgaaagtctggtgactggtcccaagaaggtggacagttgcacgaa 442
 Qy 121 gtgctagtacacggatcagttggaactgacgagagactgcaagaagaactaaata 180
 Db 443 gtgctagtacacggatcagttggaactgacgagagactgcaagaagaactaaata 502
 Qy 181 gacgtcaaaagcctgtctctggtcgcagctggaccagccctccgacagtttaaccag 240
 Db 503 gacgtcaaaagcctgtctctggtcgcagctggaccagccctccgacagtttaaccag 562
 Qy 241 tcagtgaacatgaactgaattgagtagtaggagacttctctctgctactgatggg 300
 Db 563 tcagtgaacatgaactgaattgagtagtaggagacttctctctgctactgatggg 622
 Qy 301 cagcttcatgtcaggcaaaactcctgactcctgaggtctcccaagaagaatgtactattacct 360
 Db 623 cagcttcatgtcaggcaaaactcctgactcctgaggtctcccaagaagaatgtactattacct 682
 Qy 361 gaatgcttctattctcttttttgattcttcgaaaaagaattcgaagaattgctccgtgttca 420
 Db 683 gaatgcttctattctcttttttgattcttcgaaaaagaattcgaagaattgctccgtgttca 742
 Qy 421 cctgattgacaactggagcttgccacaatgacagagatttttaatttttgagaagagt 480

Db 743 cctgattgacaactggagcttgccacaatgacagagatttttaatttttgagaagagt 802
 Qy 481 agttcagttctctgatatggagcctctcaagtgaagatatgggaaatataatttttagca 540
 Db 803 agttcagttctctgatatggagcctctcaagtgaagatatgggaaatataatttttagca 862
 Qy 541 atgatttcagagccttataatcacaggtttttcagatccagagagagtgaaattacaagt 600
 Db 863 atgatttcagagccttataatcacaggtttttcagatccagagagagtgaaattacaagt 922
 Qy 601 gaaagtgaacttcagacagatggaacttattgatatcaacacgtagtcagggcacga 660
 Db 923 gaaagtgaacttcagacagatggaacttattgatatcaacacgtagtcagggcacga 982
 Qy 661 ggtttaccatggcagctcttcagatcaagatatgccaagattcttcaaggagactcaatt 720
 Db 983 ggtttaccatggcagctcttcagatcaagatatgccaagattcttcaaggagactcaatt 1042
 Qy 721 gccaaaggaggtgcagacactttgtcgaatgctcaggggtcgaaagacggagagcctctg 780
 Db 1043 gccaaaggaggtgcagacactttgtcgaatgctcaggggtcgaaagacggagagcctctg 1102
 Qy 781 gttaggttttaagtggagcaccgagacccagctacactacagagggcacaaacatcacatg 840
 Db 1103 gttaggttttaagtggagcaccgagacccagctacactacagagggcacaaacatcacatg 1162
 Qy 841 gggaccgggtatattgaggtttacaagaacacaggtgaaagatttcccttaaatgtcgtg 900
 Db 1163 gggaccgggtatattgaggtttacaagaacacaggtgaaagatttcccttaaatgtcgtg 1222
 Qy 901 ggtacttccaatgaggttagccagtttcttccaaagaaatacaagtcattgttcgcatg 960
 Db 1223 ggtacttccaatgaggttagccagtttcttccaaagaaatacaagtcattgttcgcatg 1282
 Qy 961 cgggggctcctttccagcggccacagctgaaagaagtgtggccttcttggacagcattgc 1020
 Db 1283 cgggggctcctttccagcggccacagctgaaagaagtgtggccttcttggacagcattgc 1342
 Qy 1021 cctattactgggggaaaggagcagctcctttgtccactaccagatggtagccaca 1080
 Db 1343 cctattactgggggaaaggagcagctcctttgtccactaccagatggtagccaca 1402
 Qy 1081 ggggacgcttttgcctcttgcctgtgaggaatatgcacagaatgctgtgaggaagcat 1140
 Db 1403 ggggacgcttttgcctcttgcctgtgaggaatatgcacagaatgctgtgaggaagcat 1462
 Qy 1141 aaagacttgttggtataaagatacatgaaactcttcagagcacagcagctgaaagtccag 1200
 Db 1463 aaagacttgttggtataaagatacatgaaactcttcagagcacagcagcagctgaaagtccag 1522
 Qy 1201 caggtgctgaatcgtattctcctggccctctcctcacttccacttccaaacccctccattatt 1260
 Db 1523 caggtgctgaatcgtattctcctggccctctcctcacttccacttccaaacccctccattatt 1582
 Qy 1261 ccagttactacctaagaatttgcctcccttacaataatgttagagactgtatacgccttcca 1320
 Db 1583 ccagttactacctaagaatttgcctcccttacaataatgttagagactgtatacgccttcca 1642
 Qy 1321 ggtctccctcattgagcccaaatgagagacatcctgattcctcctgggaggttcgcacaca 1380
 Db 1643 ggtctccctcattgagcccaaatgagagacatcctgattcctcctgggaggttcgcacaca 1702
 Qy 1381 gatattcgtactcatgggtttcacatggtttgaaatcaccagggcccccacatcaggagat 1440
 Db 1703 gatattcgtactcatgggtttcacatggtttgaaatcaccagggcccccacatcaggagat 1762
 Qy 1441 gcttttaccagatgaagtctgcggacagagcatttatgctgcacagaagtgtcataaa 1500
 Db 1763 gcttttaccagatgaagtctgcggacagagcatttatgctgcacagaagtgtcataaa 1822
 Qy 1501 aaaaacatgaagcacagatattgtgaagtcttccagtgctcagctgagggagatgaacttt 1560
 Db 1823 aaaaacatgaagcacagatattgtgaagtcttccagtgctcagctgagggagatgaacttt 1882

```
QY 1561 gtgtaatggggggaactttaaatcgaaatggcttatccccacgc-----cca 1608
DB 1883 gtgtaatggggggaactttaaatcgaaatggcttatccccacgcgaatgtaattacca 1942
QY 1609 tgcctgtctctcctcctcctacacattccagctcctgtcgcagttattctcacagaagct 1668
DB 1943 tgcctgtctctcctcctcctacacattccagctcctgtcgcagttattctcacagaagct 2002
QY 1669 gccatttaccagccctgtgattttaatccacagacgactgcagccctccacagcgctac 1728
DB 2003 gccatttaccagccctgtgattttaatccacagacgactgcagccctccacagcgctac 2062
QY 1729 taccacagcagcactcagctcttcatgaactacacagcgtactatccagcccccaggt 1788
DB 2063 taccacagcagcactcagctcttcatgaactacacagcgtactatccagcccccaggt 2122
QY 1789 tcgcctaatagcttggctactctccctacagctgctaatcttagcgtgtccctccacag 1848
DB 2123 tcgcctaatagcttggctactctccctacagctgctaatcttagcgtgtccctccacag 2182
QY 1849 cctgacacgctgctcagaatcgaggcctgctcctacataactgagtttaaggaaattctt 1908
DB 2183 cctgacacgctgctcagaatcgaggcctgctcctacataactgagtttaaggaaattctt 2242
QY 1909 aacttcttccaaggttaccagtatgcacacgaggtgcttatacacacaaaatgacacag 1968
DB 2243 aacttcttccaaggttaccagtatgcacacgaggtgcttatacacacaaaatgacacag 2302
QY 1969 gccaggactaccaccaagaatgggtttgtatttaaggcccccagcagttagaacatct 2028
DB 2303 gccaggactaccaccaagaatgggtttgtatttaaggcccccagcagttagaacatct 2362
QY 2029 cagaaaagaagtttgaagaatgtatggtgatcttgaacctccagacacagaagaact 2088
DB 2363 cagaaaagaagtttgaagaatgtatggtgatcttgaacctccagacacagaagaact 2422
QY 2089 tctagcaaatcagggaagtgtctacactcaggtcgcagttatttcagcaaaccttga 2148
DB 2423 tctagcaaatcagggaagtgtctacactcaggtcgcagttatttcagcaaaccttga 2482
QY 2149 ttggacaaacggcctgtgcctattcttgggtgagtgaaaaaatttgagctagtgaag 2208
DB 2483 ttggacaaacggcctgtgcctattcttgggtgagtgaaaaaatttgagctagtgaag 2542
QY 2209 ccaaatcgtaactcacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2268
DB 2543 ccaaatcgtaactcacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2602
QY 2269 aggcatttaaaatgtgaatttgaatcagatgtctccattctccagttccagttggaat 2328
DB 2603 aggcatttaaaatgtgaatttgaatcagatgtctccattctccagttccagttggaat 2662
QY 2329 cataggtgttcttaagtcttgaatcttgataaaaaactccacagctgtctacacatctcc 2388
DB 2663 cataggtgttcttaagtcttgaatcttgataaaaaactccacagctgtctacacatctcc 2722
QY 2389 accatgaactctgttaaggagcttcttcttcttatctccgcctgttctctctctcatt 2448
DB 2723 accatgaactctgttaaggagcttcttcttcttatctccgcctgttctctctctcatt 2782
QY 2449 cccctgttcttctgcataatcatgccttcttctgaagttaattcaagcataaagatcttggaa 2508
DB 2783 cccctgttcttctgcataatcatgccttcttctgaagttaattcaagcataaagatcttggaa 2842
QY 2509 aataaaatcaaatctttaggagaaagaataaaattgt 2545
DB 2843 aataaaatcaaatctttaggagaaagaataaaatttt 2879
```

RESULT 3
AAF98702
ID AAF98702 standard; DNA; 2731 BP.

```
XX AAF98702;  
XX AC  
XX DE  
DT 02-JUL-2001 (first entry)  
XX Human ovarian cancer cell expressed sequence 10802.  
XX DE  
XX KW Human; ovarian cancer; identification; detection; characterisation;  
KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.  
XX OS Homo sapiens.  
XX PN W0200118542-A2.  
XX PD 15-MAR-2001.  
XX PF 01-SEP-2000; 2000WO-US24199.  
XX PR 03-SEP-1999; 99US-0152547.  
PR 16-MAR-2000; 2000US-0190347.  
PR 21-MAR-2000; 2000US-0191321.  
PR 31-MAY-2000; 2000US-0208382.  
PR 20-JUL-2000; 2000US-0220467.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Lee J, Thompshe P, Lillie J;  
XX WPI; 2001-211428/21.  
XX DR  
XX PT Detection, assessment, prevention and therapy of ovarian cancer,  
PT comprises detecting changes in the expression of a variety of markers -  
XX  
XX PS Claim 1; Page 1004-1005; 1198pp; English.  
XX CC The present invention describes a method for assessing whether a patient  
CC is afflicted with ovarian cancer by comparing: (1) the expression of a  
CC marker (1) (see AAF98594 to AAF98730), in a patient sample; and (2) the  
CC normal level of expression of (1) in a control non-ovarian cancer  
CC sample, where a significant difference between the level of expression  
CC in (a) and (b) is an indication that the patient is afflicted with  
CC ovarian cancer. (1) have cytostatic activities and can be used in  
CC antisense gene therapy. The method, compositions and kits from the  
CC present invention can be used for: (1) assessing and treating ovarian  
CC cancer; (2) making isolated hybridoma, which produces an antibody useful  
CC for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a  
CC patient. AAF98573 to AAF98593 represent human kinase marker primers and  
CC probes which are used in the exemplification of the present invention.  
XX SQ Sequence 2731 BP; 825 A; 562 C; 549 G; 795 T; 0 other;  
  
Query Match 65.8%; Score 2360; DB 22; Length 2731;  
Best Local Similarity 90.4%; Pred. No. 0;  
Matches 2659; Conservative 0; Mismatches 10; Indels 273; Gaps 2;  
  
QY 656 caccaggtttaccatgacgtcttcacatcaagatattgcaagattcttcaagactca 715  
DB 16 cgcgaggtttaccatgacgtcttcacatcaagatattgcaagattcttcaagactca 75  
QY 716 atattgccaaaggagggtgcagcactttgtctgaatgctcaggttcgaaggacggagaag 775  
DB 76 atattgccaaaggagggtgcagcactttgtctgaatgctcaggttcgaaggacggagaag 135  
QY 776 ctctggttaggtttgaagtgcagcaccgagaccctagcactacagaggcacaacaatc 835  
DB 136 ctctggttaggtttgaagtgcagcaccgagaccctagcactacagaggcacaacaatc 195  
QY 836 acatggggaccggttatattgagttttacaagaacaggtggaagattccttaaatg 895  
DB 196 acatggggaccggttatattgagttttacaagaacaggtggaagattccttaaatg 255  
QY 896 ctggtggtaactccaatgaggtagccagtttctctccaaggaaaatccaagtcattgttc 955
```

Db 256 ||||| ctggtgtaactccaatgagtgagccagttctctccaaggaaaaatcaagtcatgtcc 315
 QY 956 gcatcgggggccctttcacagccacacagctgaagaagtgggtggtctctttggacagc 1015
 Db 316 gcatcgggggccctttcacagccacacagctgaagaagcgggtggtctctttggacagc 375
 QY 1016 attgcccattactgggggaaaggaaaggaatcctctttgtcacctaccagatggtaggc 1075
 Db 376 attgcccattactgggggaaaggaaaggaatcctctttgtcacctaccagatggtaggc 435
 QY 1076 caacaggggacgttttgcctctcttgcctgtgaggaatgacacagaatgcgttagga 1135
 Db 436 caacaggggacgttttgcctctcttgcctgtgaggaatgacacagaatgcgttagga 495
 QY 1136 agcataaagactgttgggttaaagatacaattgaactcttcaggagacagcagctgaag 1195
 Db 496 agcataaagactgttgggttaaagatacaattgaactcttcaggagacagcagctgaag 555
 QY 1196 ttacagcaggtgctgaatcgattctctcgccctctctcattccacttccaaacctccca 1255
 Db 556 ttacagcaggtgctgaatcgattctctcgccctctctcattccacttccaaacctccca 615
 QY 1256 ttattccagtaactacagcaatttgcctccctcaaaatgtagagactgtatacgc 1315
 Db 616 ttattccagtaactacagcaatttgcctccctcaaaatgtagagactgtatacgc 675
 QY 1316 ttccgaggtctccctatgcagccacaaatgagacatcctcggttccctggggaggtcg 1375
 Db 676 ttccgaggtctccctatgcagccacaaatgagacatcctcggttccctggggaggtcg 735
 QY 1376 ccacagatattcgtaactatgggtttcacatggttttgaatcaccaggggcccccacag 1435
 Db 736 ccacagatattcgtaactatgggtttcacatggttttgaatcaccaggggcccccacag 795
 QY 1436 gagatgcctttatccagatgaagctgcggacagagcatttatggctgacagaagtgtc 1495
 Db 796 gagatgcctttatccagatgaagctgcggacagagcatttatggctgacagaagtgtc 855
 QY 1496 ataaaaaaacatgaaggacagatgttgaagctcttccagtttcagctgcaggagatga 1555
 Db 856 ataaaaaaacatgaaggacagatgttgaagctcttccagtttcagctgcaggagatga 915
 QY 1556 actttgttaattggggggacgtttaaatcgaaatggcttatccccacgcg----- 1605
 Db 916 actttgttaattggggggacgtttaaatcgaaatggcttatccccacgcgccaatgtaagt 975
 QY 1606 --ccatgctgtctcctcctctacacatttccagctcctgtgcagttattcctacag 1663
 Db 976 taccatgctgtcctcctcctcctacacatttccagctcctgtgcagttattcctacag 1035
 QY 1664 aagctgccatttaccagccctctgtgatgttgaatccagagcactgcagccctccacag 1723
 Db 1036 aagctgccatttaccagccctctgtgatgttgaatccagagcactgcagccctccacag 1095
 QY 1724 cgtactaccagcagcactcagctcttcatgaactacacagcgtactatccccagcccc 1783
 Db 1096 cgtactaccagcagcactcagctcttcatgaattacacagcgtactatcccc----- 1148
 QY 1784 caggttcgcctaagtcttggctacttccctacagctgctaatttagcgtgtccctc 1843
 Db 1149 ----- 1148
 QY 1844 cacagctggcacggtggtcagaatgcagggcctgacctacatactggagttaaagaaa 1903
 Db 1149 ----- 1148
 QY 1904 ttcttaacttctccaaggttaccagtatgcaaccaggaggtggacttatcacacaaaatg 1963
 Db 1149 ----- 1148
 QY 1964 accagccaggactctaaccaagaatgggtttgtatttaaggggcccccagcagttagaac 2023

Db 1149 ----- 1148
 QY 2024 atcctcagaaaaagaagtgtttgaaagatgtatggtgatcttgaacctccagacacaaga 2083
 Db 1149 ----- agtgtttgaaagatgtatggtgatcttgaacctccagacacaaga 1194
 QY 2084 aaacttttagcaaatccaggggagtttgtacaactcagggtcagctattttcagcaaaa 2143
 Db 1195 aaacttttagcaaatccaggggagtttgtacaactcagggtcagctattttcagcaaaa 1254
 QY 2144 ctgagattggaacaacggcgtgtgccttatcttttgggtgagtgaaaaaatttagctag 2203
 Db 1255 ctgagattggaacaacggcgtgtgccttatcttttgggtgagtgaaaaaatttagccag 1314
 QY 2204 tgaagccaaatcgtaaacttacagcaagcagcatgcagcatcacctggctctttgctgattg 2263
 Db 1315 tgaagccaaatcgtaaacttacagcaagcagcatgcagcatcacctggctctttgctgattg 1374
 QY 2264 caaataggcattttaaagtgtgaatttggaaatcagatgtctccatttacttccagttaaagt 2323
 Db 1375 caaataggcattttaaagtgtgaatttggaaatcagatgtctccatttacttccagttaaagt 1434
 QY 2324 ggcatacaggtgttttccctaagitttaagctcttgataaaaaactccaccaggtgttaacca 2383
 Db 1435 ggcatacaggtgttttccctaagittttaaagctcttgataaaaaactccaccaggtgttaacca 1494
 QY 2384 tctccaccatgaactctgttaaaggaaagcttcaatttttgtatattcccgctctttctctt 2443
 Db 1495 tctccaccatgaactctgttaaaggaaagcttcaatttttgcgtatattcccgctctttctctt 1554
 QY 2444 cattccctgtcttctgcataatcatcgtccttctgttaagttaatttcaagcataagaattt 2503
 Db 1555 cattccctgtcttctgcataatcatcgtccttctgttaagttaatttcaagcataagaattt 1614
 QY 2504 ggaataataaaatcacaaatotttagggagaaagaataaaattgttatatttccagctctcttg 2563
 Db 1615 ggaataataaaatcacaaatotttagggagaaagaataaaattgttatatttccagctctcttg 1674
 QY 2564 gccatgatgatacttatgattaaaaacaaatlaaaattttaaaccacctgaagataaatt 2623
 Db 1675 gccatgatgatacttatgattaaaaacaaatlaaaattttaaaccacctgaagataaatt 1734
 QY 2624 agaagaaattgtcacccctccacaaacatacaaaagtttaaaagtttggattctttctctc 2683
 Db 1735 agaagaaattgtcacccctccacaaacatacaaaagtttaaaagtttggattctttctc 1794
 QY 2684 acagatatacagttgataataatgaattaggggcccaaaatgcaaaacgaaaaatgaagca 2743
 Db 1795 agcaggtatcagttgtataataatgaattaggggcccaaaatgcaaaacgaaaaatgaagca 1854
 QY 2744 gctacatgtagttagtaaatcttagtttgaactgtaattgaattgtggtcttcatagt 2803
 Db 1855 gctacatgtagttagtaaatcttagtttgaactgtaattgaattgtggtcttcatagt 1914
 QY 2804 attattttatattgtaacttttttccattatgtatggtttggactttaataagagaaaattcc 2863
 Db 1915 attattttatattgtaacttttttccattatgtatggtttggactttaataagagaaaattcc 1974
 QY 2864 atagtttttaataatccagagagagacaatttgaacaggtattcttagaaaaacaataca 2923
 Db 1975 atagtttttaataatccagagagagacaatttgaacaggtattcttagaaaaacaataca 2034
 QY 2924 ctaactgaacagagagagaaatgcttatataattatgatagccttaaaccttttccctcta 2983
 Db 2035 ctaactgaacagagagagaaatgcttatataattatgatagccttaaaccttttccctcta 2094
 QY 2984 atgccttaactgtcaaaataattataaccttttaagcataggactatagtcagcatgcta 3043
 Db 2095 atgccttaactgtcaaaataattataaccttttaagcataggactatagtcagcatgcta 2154
 QY 3044 gactgagaggtaaacactgatcaatttagaacaggtactgatgctgctcaggtgttaaacac 3103
 Db 2155 gactgagaggtaaacactgatgcaatttagaacaggtactgatgctgctcaggtgttaaacac 2214

QY 3104 tatgttttagctgtgtttatgctataaaagtgcgaatattagacactagctagctactgctgc 3163
|||||
Db 2215 tatgttttagctgtgtttatgctataaaagtgcgaatattagacactagctagctactgctgc 2274
QY 3164 ctcattgtaactccaaagaacagaggatttcattaaagtgcattgaatggtgataattctct 3223
Db 2275 ctcattgtaactccaaagaacagaggatttcattaaagtgcattgaatggtgataattctct 2334
QY 3224 agttactcatattgctcttcttgtaagtgcgaatgcggtgcagatttatgagctgctat 3283
|||||
Db 2335 agttactcatattgctcttcttgtaagtgcgaatgcggtgcagatttatgagctgctat 2394
QY 3284 ttttattttctgtgctattactttaacaccttaagggagagcaaacatttctcttcca 3343
Db 2395 ttttattttctgtgctattactttaacaccttaagggagagcaaacatttctcttcca 2454
QY 3344 gctgactggaatgccccttttaactgcaataggaagaaacaaaggtttgtgtaaa 3403
Db 2455 gctgactggaatgccccttttaactgcaataggaagaaacaaaggtttgtgtaaa 2514
QY 3404 attggtgataactgacacaaagctctgaaagactttaagataggcagtaagtcttactacaat 3463
Db 2515 attggtgataactgacacaaagctctgaaagactttaagataggcagtaagtcttactacaat 2574
QY 3464 tgcccaaaagctgccccttttaactgcaataggaagaaacaaaggtttgtgtaaa 3523
Db 2575 tgcccaaaagctgccccttttaactgcaataggaagaaacaaaggtttgtgtaaa 2634
QY 3524 actactgagttttgttagaggttaaacatttgataataaaactgctgtttaaactcctcaaa 3583
Db 2635 actactgagttttgttagaggttaaacatttgataataaaactgctgtttaaactcctcaaa 2694
QY 3584 aa 3585
Db 2695 aa 2696
RESULT 4
AAH33714
ID AAH33714 standard; cDNA; 2468 BP.
XX
AC AAH33714;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:770.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
DR P-PSDB; AAG74283.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
PS Claim 1; Page 2745-2746; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 2468 BP; 753 A; 508 C; 466 G; 733 T; 8 other;

Query Match 64.6%; Score 2316.8; DB 22; Length 2468;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2432; Conservative 8; Mismatches 0; Indels 110; Gaps 1;

QY 1036 agggaggagctcctttgtcacctaccagatggttagcccaacagggcgcttttgc 1095
Db 2 aaggaggagctcctttgtcacctaccagatggttagcccaacagggcgcttttgc 61
QY 1096 ctcttgcctgtgaggaatatgcagaatgcgtttgagaagcataaagacttgttgggt 1155
Db 62 ctcttgcctgtgaggaatatgcagaatgcgtttgagaagcataaagacttgttgggt 121
QY 1156 aaagatacattgaactcttcaggagcacagcagctgaagtgcagaggttgcgaatcga 1215
Db 122 aaagatacattgaactcttcaggagcacagcagctgaagtgcagaggttgcgaatcga 181
QY 1216 ttctctcgccctctcattccacttccaaacccctcccaatttccagttactactcag 1275
Db 182 ttctctcgccctctcattccacttccaaacccctcccaatttccagttactactcag 241
QY 1276 caatttgcctccctacaaatgttagagctgtatagccttcgaggtctcctctatgca 1335
Db 242 caatttgcctccctacaaatgttagagctgtatagccttcgaggtctcctctatgca 301
QY 1336 gccacaattgagacacatcctggatttctctggggaggttcgccacagatatctactcat 1395
Db 302 gccacaattgagacacatcctggttctctggggaggttcgccacagatatctactcat 361
QY 1396 ggggttcacatggttttgatcacccggcgcccatccatccagagatgctttatccagatg 1455
Db 362 ggggttcacatggttttgatcacccggcgcccatccatccagagatgctttatccagatg 421
QY 1456 aagctcgagacagacattttatgctgcacagaagtgcataaaaaaacatgaaggac 1515
Db 422 aagctcgagacagacattttatgctgcacagaagtgcataaaaaaacatgaaggac 481
QY 1516 agatatgttgaagtcttccagtttccagctgagagatgaactttgtttaatggggggc 1575
Db 482 agatatgttgaagtcttccagtttccagctgagagatgaactttgtttaatggggggc 541
QY 1576 actttaaatcgaaatggcttatcccccacgcgatgctgtctctccctccctcacattt 1635
Db 542 actttaaatcgaaatggcttatcccccacgcgatgctgtctctccctccctcacattt 601
QY 1636 ccagctctctgctgcaatttctctacagaagctgcatttacagcctctgtgatttg 1695
Db 602 ccagctctctgctgcaatttctctacagaagctgcatttacagcctctgtgatttg 661
QY 1696 aatccacagcactgcagccctccacagcgtactaccacagcaggcactcagcttctcatg 1755

[illegible]

Db	1632	tgtaattgaataattggtgcctcaatgtaattattttatattgtacttttttttctattatgta	1691
Qy	2836	tgttttgaccttaataagagaaattccatagtttttaataatccagaagtgcagacaatt	2895
Db	1692	tgttttgaccttaataagagaaattccatagtttttaataatccagaagtgcagacaatt	1751
Qy	2896	tgaacagtgtattcttagaaaaacaatacacataactgaacagaagtgaatgcttatatat	2955
Db	1752	tgaacagtgtattcttagaaaaacaatacacataactgaacagaagtgaatgcttatatat	1811
Qy	2956	tatgatagccttaaacctttttcccttaatagccttaactgaactcaaatataaactttat	3015
Db	1812	tatgatagccttaaacctttttcccttaatagccttaactgaactcaaatataaactttt	1871
Qy	3016	aaagcataggactatagtcagcatgctagacgtgagaggttaaacactgatgcgaattagaac	3075
Db	1872	aaagcataggactatagtcagcatgctagacgtgagaggttaaacactgatgcgaattagaac	1931
Qy	3076	aggtactgatgctgcagtggtttaaacactatgttttagctgtgtttatgctataaaagtgc	3135
Db	1932	aggtactgatgctgcagtggtttaaacactatgttttagctgtgtttatgctataaaagtgc	1991
Qy	3136	aatattagaactagctagtaactgctgcctatgtaactccaaagaaaacaggatttcatt	3195
Db	1992	aatattagaactagctagtaactgctgcctatgtaactccaaagaaaacaggatttcatt	2051
Qy	3196	taagtgcattgaatggatattctctaagttactcatatgtctcttgcttgaatgca	3255
Db	2052	taagtgcattgaatggatattctctaagttactcatatgtctcttgcttgaatgca	2111
Qy	3256	atgcctgctcagatttataggctgctattttattttctgtgcattactttaacacctta	3315
Db	2112	atgcctgctcagatttataggctgctattttattttctgtgcattactttaacacctta	2171
Qy	3316	aaggagagaagcaaacatttccttcttcagctgaactggcaatggccctttaactgcgaatag	3375
Db	2172	aaggagagaagcaaacatttccttcttcagctgaactggcaatggccctttaactgcgaatag	2231
Qy	3376	gaagaaaaaaaagaggttgtgtgaaaattggtgataactggcacttaagatcgaaaaag	3435
Db	2232	gaagaaaaaaaagaggttgtgtgaaaattggtgataactggcacttaagatcgaaaaag	2291
Qy	3436	aaattctgtatactttagcttaagtgcgcaaaagctgccaaagctgccaaagctgaaagactt	3495
Db	2292	aaattctgtatactttagcttaagtgcgcaaaagctgccaaagctgccaaagctgaaagactt	2351
Qy	3496	taagataggcagtaatgcttactacaataactagtttttgtaaggttaaacatttgat	3555
Db	2352	taagataggcagtaatgcttactacaataactagtttttgtaaggttaaacatttgat	2411
Qy	3556	aataaaactgcctgtttaactctcaaaaaa	3585
Db	2412	aataaaactgcctgtttaactctcaaaaaa	2441
RESULT	5		
AC	AAC77930		
ID	AAC77930 standard; cDNA; 2261 BP.		
XX			
AC	AAC77930;		
XX			
DT	08-FEB-2001 (first entry)		
XX			
DE	Human cancer associated gene sequence SEQ ID NO:324.		
XX			
KW	Human; cancer associated gene; cancer antigen; detection; cancer;		
KW	diagnosis; cytostatic; proliferative; cancer antigen; immunomodulator;		
KW	antidiabetic; antiasthmatic; antirheumatic; antithrombotic; antiviral;		
KW	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;		
KW	dermatological; neuroprotective; thrombolytic; coagulant; neotropic;		
KW	vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;		
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;		
KW	allergic reaction; graft versus host disease; organ rejection;		

RECHT M

RESULT
ABC7793

ID AAC77930 standard: cDNA: 2261 BP.

XX XX

AC AAC77930;

XX

DT 08-FEB-2001 (first entry)

xx
DE Human cancer associated gene sequence SEQ ID NO:324.

XX

KW Human; cancer associated gene; cancer antigen; cancer detection; cancer;

KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;

KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antitumor

KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardi

KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic

KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation

KW Immune disorder; haematopoietic cell disorder; autoimmune disorder

KW allergic reaction; graft versus host disease; organ rejection;

QY 2442 ttcatctccctgtcttctgcatataatcctgctcttcttctgtaagtaattcaagcataagatc 2501
 Db 1120 ttcatctccctgtcttctgcatataatcctgctcttcttctgtaagtaattcaagcataagatc 1179
 QY 2502 ttggaataataaaatcacaatcttaggagaagaataaaaattgtattttccagctctct 2561
 Db 1180 ttggaataataaaatcacaatcttaggagaagaataaaaattgtattttccagctctct 1239
 QY 2562 tggccatgatgatctcttattgattataaaacaaataataattttataaacacctggaagataaa 2621
 Db 1240 tggccatgatgatctcttattgattataaaacaaataataattttataaacacctggaagataaa 1299
 QY 2622 ttagaagaattgtgcacctccacaaacatacacaagtgttttaaaagtgttgatcttttc 2681
 Db 1300 ttagaagaattgtgcacctccacaaacatacacaagtgttttaaaagtgttgatcttttc 1359
 QY 2682 tcagcaggtatcagttgtaataatgaattaggggccaaaatgcaaaacgaaaatgaag 2741
 Db 1360 tcagcaggtatcagttgtaataatgaattaggggccaaaatgcaaaacgaaaatgaag 1419
 QY 2742 cagctcacatgtagtttagtaattcttagttgaactgttaattgaattgtggtctcatat 2801
 Db 1420 cagctcacatgtagtttagtaattcttagttgaactgttaattgaattgtggtctcatat 1479
 QY 2802 gtattatttatattgtactttttcttcttattatgtatggtttggacctttaataagagaatt 2861
 Db 1480 gtattatttatattgtactttttcttcttattatgtatggtttggacctttaataagagaatt 1539
 QY 2862 coatagttttaataatccagagtgagacaatttgaaactgttaattgaattgtggtctcatat 2921
 Db 1540 coatagttttaataatccagagtgagacaatttgaaactgttaattgaattgtggtctcatat 1599
 QY 2922 cactaactgaacagagtgaaatgcttatataattatgatagccttaaacctttttctctc 2981
 Db 1600 cactaactgaacagagtgaaatgcttatataattatgatagccttaaacctttttctctc 1659
 QY 2982 taatgccttaactgcataataattataaccttttaaacctataggactatagtcagcatg 3041
 Db 1660 taatgccttaactgcataataattataaccttttaaacctataggactatagtcagcatg 1719
 QY 3042 tagactgagagtgaaactgtatgaatttagaacagagtgactgtctgctgagtttaac 3101
 Db 1720 tagactgagagtgaaactgtatgaatttagaacagagtgactgtctgctgagtttaac 1779
 QY 3102 actatgttttagctgtgtttatgctataaaagtgcataattatagacactgactgtactgct 3161
 Db 1780 actatgttttagctgtgtttatgctataaaagtgcataattatagacactgactgtactgct 1839
 QY 3162 gctcatgttaactcaagagaacaggatttcatatgaattgaattgaattgataatttt 3221
 Db 1840 gctcatgttaactcaagagaacaggatttcatatgaattgaattgaattgataatttt 1899
 QY 3222 ctaagtactcatattgtcttctgctgaatgcaatgcctgacagatttatgagctgct 3281
 Db 1900 ctaagtactcatattgtcttctgctgaatgcaatgcctgacagatttatgagctgct 1959
 QY 3282 atttttttttctgcttactttaaacaccttaaaaggagagaacaaattcttctcttct 3341
 Db 1960 atttttttttctgcttactttaaacaccttaaaaggagagaacaaattcttctcttct 2019
 QY 3342 cagctgactggaactggccctttaaactgcaatagagaagaaaaaaagggtttgtgga 3401
 Db 2020 cagctgactggaactggccctttaaactgcaatagagaagaaaaaaagggtttgtgga 2079
 QY 3402 aaattggtgataactggcacttaagatgaaagaaattctgtactactgtgactgactaa 3461
 Db 2080 aaattggtgataactggcacttaagatgaaagaaattctgtactactgtgactgactaa 2139
 QY 3462 gatgcccagaagctgccaaagctctgaaagactttaagatagcgagtaattgttactaca 3521
 Db 2140 gatgcccagaagctgccaaagctctgaaagactttaagatagcgagtaattgttactaca 2199
 QY 3522 atactactgagtttttctgtagagtttaacatttgataataaaactgctgtttaactcaaa 3581

Db 2200 atactactgagtttttctgtagagtttaacattgataataaaactgctgtttaactcaaa 2259
 QY 3582 aa 3583
 Db 2260 aa 2261
 RESULT 6
 AAH14740
 ID AAH14740 standard; cDNA: 1624 BP.
 XX AAH14740;
 AC AAH14740;
 XX 26-JUN-2001 (first entry)
 XX Human cDNA sequence SEQ ID NO:12485.
 DE Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8; SEQ ID 12485; 2537pp + CD ROM; English.
 PS The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX Sequence 1624 BP; 498 A; 320 C; 288 G; 518 T; 0 other;

Query Match									
Best Local Similarity 37.0%; Score 1324.8; DB 22; Length 1624;									
Matches 1329; Conservative 0; Mismatches 7; Indels 0; Gaps 0;									
QY	2038	agtggttgaagaagatgtatgtgatcttgaaacctccagacacaagaataacttctagcaaa	2097						
DB	289	agtggttgaagaagtatgtgacttgaaacctccagacacaagaataacttctagcaaa	348						
QY	2098	tccaggggaagttgtctacactcaggctgcagtattttccagcaaatctgattggacaaa	2157						
DB	349	tccaggggaagttgtctacactcaggctgcagtattttccagcaaatctgattggacaaa	408						
QY	2158	cgggcctgtgccttatcttttggtyggagtgaaaaatttgagctagtgaaagccaaatcgt	2217						
DB	409	cgggcctgtgccttatcttttggtyggagtgaaaaatttgagccagtgaaagccaaatcgt	468						
QY	2218	aacttacagaagaagcagatgcagatcacctgcgtctctttgctgattgcaaatagcattta	2277						
DB	469	aacttacagaagaagcagatgcagatcacctgcgtctctttgctgattgcaaatagcattta	528						
QY	2278	aaatgtgaatttggaaatcagatgtctccattcaactccagttcaagtgggcatcagtggt	2337						
DB	529	aaatgtgaatttggaaatcagatgtctccattcaactccagttcaagtgggcatcagtggt	588						
QY	2338	ttcctaagttttaaagtccttggtataaaactccaccagtgctacacatccaccatgaac	2397						
DB	589	ttcctaagttttaaagtccttggtataaaactccaccagtgctacacatccaccatgaac	648						
QY	2398	tcgtttaagaagcttcatttttggatatattccgcgtcttttctcttcaatttccctgtctt	2457						
DB	649	tcgtttaagaagcttcatttttggatatattccgcgtcttttctcttcaatttccctgtctt	708						
QY	2458	ctgcataatcatgctctctgtctaagttaattccaagcataagatcttggaaataaaaaac	2517						
DB	709	ctgcataatcatgctctctgtctaagttaattccaagcataagatcttggaaataaaaaac	768						
QY	2518	acaactctaggagaaagataaaattgttatttccacagtcctcttggccatgatgatc	2577						
DB	769	acaactctaggagaaagataaaattgttatttccacagtcctcttggccatgatgatc	828						
QY	2578	ttatgattaaaaacaaattaaaattttaaaccacctgaagataaattagaagaaaattgtgc	2637						
DB	829	ttatgattaaaaacaaattaaaattttaaaccacctgaagataaattagaagaaaattgtgc	888						
QY	2638	accctccacaaaacatacaagttttaaagtttggatctttttctcagcaggtatcagtt	2697						
DB	889	accctccacaaaacatacaagttttaaagtttggatctttttctcagcaggtatcagtt	948						
QY	2698	gtaaataatgaattaggggccaaaatgcaaaacgaaaaatgaagcagctacatgattta	2757						
DB	949	gtaaataatgaattaggggccaaaatgcaaaacgaaaaatgaagcagctacatgattta	1008						
QY	2758	gtaatttctagtttgaactgttaattgtaatatgttggtcttcattgtattttttatttg	2817						
DB	1009	gtaatttctagtttgaactgttaattgtaatatgttggtcttcattgtattttttatttg	1068						
QY	2818	tacttttttcaattatgtatggtttggactttaataagagaaaattccatagttttaatat	2877						
DB	1069	tacttttttcaattatgtatggtttggactttaataagagaaaattccatagttttaatat	1128						
QY	2878	cccagaagtgagacaaatttgaacagtgattctctagaaaaacaatacactaacctgaacgaa	2937						
DB	1129	cccagaagtgagacaaatttgaacagtgattctctagaaaaacaatacactaacctgaacgaa	1188						
QY	2938	gtgaatgcttatatatattatgatagccttaaaacctttttctcctaaatgccttaactgtc	2997						
DB	1189	gtgaatgcttatatatattatgatagccttaaaacctttttctcctaaatgccttaactgtc	1248						
QY	2998	aaataataaaccttttaagacatagactatatgtcagcatgtctagactgagaggtataa	3057						
DB	1249	aaataataaaccttttaagacatagactatatgtcagcatgtctagactgagaggtataa	1308						

QY	3058	cactgatgcgaattgagacaggtactgatgtctgcagtgctttaaacactatgttttagctgtg	3117
Db	1309	cactgatgcgaattgagacaggtactgatgtctgcagtgctttaaacactatgttttagctgtg	1368
QY	3118	tttagctataaaagtgcgaattatagacactagctagtactgctgcctcatgtaaactcca	3177
Db	1369	tttagctataaaagtgcgaattatagacactagctagtactgctgcctcatgtaaactcca	1428
QY	3178	aagaaacagagatttcatttaagtgcattggaatgggatatctctcctaagttactcatatt	3237
Db	1429	aagaaacagagatttcatttaagtgcattggaatgggatatctctcctaagttactcatatt	1488
QY	3238	gtcctttgcttgaatgcaatgcgcgtgcagatttatgaggctgctattttattttctgtg	3297
Db	1489	gtcctttgcttgaatgcaatgcgcgtgcagatttatgaggctgctattttattttctgtg	1548
QY	3298	cattactttaacacctaaagggaaggaacaaacatttccttcagctgaactggcaatg	3357
Db	1549	cattactttaacacctaaagggaaggaacaaacatttccttcagctgaactggcaatg	1608
QY	3358	gcccttttaactgccaat	3373
Db	1609	gcccttttaactgccaat	1624
RESULT 7			
AAC59113			
ID	AAC59113 standard; cDNA; 1585 BP.		
AC	AAC59113;		
DT	02-FEB-2001 (first entry)		
DE	Human secreted protein coding sequence SEQ ID NO: 16.		
DE	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;		
KW	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;		
KW	vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;		
KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;		
KW	neurological disease; infection; human; secreted protein; ss.		
OS	Homo sapiens.		
PN	WO2000055177-A2.		
PD	21-SEP-2000.		
PF	09-MAR-2000; 2000WO-US06058.		
PR	12-MAR-1999; 99US-0124145.		
PR	03-SEP-1999; 99US-0168654.		
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Rosen CA, Ruben SM, Komatsoulis G;		
DR	WPI; 2000-638177/61.		
DR	P-PSDB; AAB28017.		
PT	Novel nucleic acids encoding 49 human secreted proteins useful for		
PT	treating cancers, hyperproliferative disorders, inflammatory disorders,		
PT	neurological disorders and cardiovascular disorders -		
PS	Claim 1; Page 320; 389pp; English.		
CC	The invention relates to the isolation of genes AAC59108-A59156 encoding		
CC	49 human secreted proteins AAB28012-B28060. The genes can be used to		
CC	generate fusion proteins by linking to the gene for the human		
CC	immunoglobulin G Fc portion (SEQID1) for increasing the stability of		
CC	the fusion protein as compared to the human protein only. The genes and		
CC	proteins are useful for preventing, ameliorating or treating medical		
CC	conditions, e.g. by protein or gene therapy. The genes are isolated		
CC	from a range of human tissues disclosed in the specification The		

CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer, and other cancers of the adrenal gland, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX Sequence 1585 BP; 385 A; 353 C; 345 G; 502 T; 0 other;

Query Match 16.4%; Score 587.4; DB 21; Length 1585;
Best Local Similarity 99.3%; Pred. No. 5.9e-141;
Matches 611; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 cttttggatcactgctggggccaccgggcccagctaggctcgatgagaaggagtgg 60
DB 155 cttttggatcactgctggggccaccgggcccagctaggctcgatgagaaggagtgg 214
QY 61 atctctgttctggaagtctggtatctggcccaacaagaagtgaggacgttgacgaa 120
DB 215 atctctgttctggaagtctggtatctggcccaacaagaagtgaggacgttgacgaa 274
QY 121 gtctagttagaccgcatcagttggaactgacggagactgcaaaagaagaactaaata 180
DB 275 --tctagttagaccgcatcagttggaactgacggagactgcaaaagaagaactaaata 332
QY 181 gacgtcgaaagcctctcctcgctgcgagctggaccagccctccgacagtttaaccag 240
DB 333 gacgtcgaaagcctctcctcgctgcgagctggaccagccctccgacagtttaaccag 392
QY 241 tcaagttagcaatgaactgaattgagtaggagacttcttctgtctgtactgat-gg 299
DB 393 tcaagttagcaatgaactgaattgagtaggagacttcttctgtctgtactgatggg 452
QY 300 gcaagcttcagtcaggcaaatcctgcatcctgaggtctccaaagaagaatgtactattacc 359
DB 453 gcaagcttcagtcaggcaaatcctgcatcctgaggtctccaaagaagaatgtactattacc 512
QY 360 tgaatgcttctattctctttttgatcttcgaaagaattcaagaagaatgtgctcctggttc 419
DB 513 tgaatgcttctattctctttttgatcttcgaaagaattcaagaagaatgtgctcctggttc 572
QY 420 acctgatattgacaaactggacttgcacaaatgacagagatttttaattttgagaagag 479
DB 573 acctgatattgacaaactggacttgcacaaatgacagagatttttaattttgagaagag 632
QY 480 tagttcagttctcgatagagcctctcaagttgaagatattggaataataatttttagc 539
DB 633 tagttcagttctcgatagagcctctcaagttgaagatattggaataataatttttagc 692
QY 540 aatgatttcagagccttataataacacaggttttcagatccagagagagtgaattacaagt 599
DB 693 aatgatttcagagccttataataacacaggttttcagatccagagagagtgaattacaagt 752
QY 600 tgaagtggaaacttg 614
DB 753 tgaagtggaaacttg 767

RESULT 8
AAZ16154
ID AAZ16154 standard; cDNA; 740 BP.
XX
XX AAZ16154;
XX AC
XX AT
XX 12-OCT-1999 (first entry)
XX Human gene expression product cDNA sequence SEQ ID NO:3624.

XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX Homo sapiens.
XX WO9938972-A2.
XX 05-AUG-1999.
XX 28-JAN-1999; 99WO-US01619.
XX 03-APR-1998; 98US-0080666.
XX 28-JAN-1998; 98US-0072910.
XX 24-FEB-1998; 98US-0075954.
XX 31-MAR-1998; 98US-0080114.
XX 03-APR-1998; 98US-0080515.
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX WPI; 1999-494092/41.
XX Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX Claim 1; Page 1729; 2479pp; English.
XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ17799. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17799. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX Sequence 740 BP; 188 A; 163 C; 183 G; 193 T; 13 other;

Query Match 15.5%; Score 555.6; DB 20; Length 740;
Best Local Similarity 98.5%; Pred. No. 6.4e-133;
Matches 591; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
QY 1 cttttggatcactgctggggccaccgggcccagctaggctcgatgagaaggagtgg 60
DB 143 cttttggatcactgctggggccaccgggcccagctaggctcgatgagaaggagtgg 202
QY 61 atctctgttctggaagtctggtatctggcccaacaagaagtgaggacaggttgacgaa 120
DB 203 atctctgttctggaagtctggtatctggcccaacaagaagtgaggacaggttgacgaa 262
QY 121 gtctagttagaccgcatcagttggaactgacggagactgcaaaagaagaactaaata 180

|||||
Db 263 ggcgtagtagaccgagtcagttgaaactgacgagagactcaaaagagaataaata 322
QY 181 gacgtcgaaagcctgtctctcggtcgagcagctggcccaagccctccgacagtttaaccag 240
Db 323 gacgtcgaaagcctgtctctcggtcgagcagctggcccaagccctccgacagtttaaccag 382
QY 241 tcagtgagcaatgaactgaattgagtaggagacttctctgtctgtactgatggg 300
Db 383 tcagtgagcaatgaactgaattgagtaggagacttctctgtctgtactgatggg 442
QY 301 cagcttcattgtagcagcaatctgcatctgaggtctccaaagaagaatgtactattac 360
Db 443 cagcttcattgtagcagcaatctgcatctgaggtctccaaagaagaatgtactattac 502
QY 361 gaatctctattctcttttggatttcgaaagaatcaagaagaatgtgacctgggtca 420
Db 503 gaatctctattctcttttggatttcgaaagaatcaagaagaatgtgacctgggtca 562
QY 421 cctgatattgacaaactgagcttgcacaaatgaca-gagtatttaattttgagaagag 479
Db 563 cctgatattgacaaactgagcttgcacaaatgaca-gagtatttaattttgagaagag 622
QY 480 tagttcagctctcgatggagcctctcaagttgaagatatggggaatatattttagc 539
Db 623 tagttcagctctcgatggagcctctcaagttgaagatatggggaatatattttagc 682
QY 540 aatgatttcagagccttataatcacaggttttcagatccagagagaggaattacaagt 599
Db 683 aatgatttcaga-ncttataatcacagg-tttcagatccagagagaggtgatttcaagt 740

RESULT 9
AAZ16155
ID AAZ16155 standard; cDNA; 745 BP.
XX
AC AAZ16155;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:3625.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN W0938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
DR WPI; 1999-494092/41.
XX
PT Novel human genes and their expression products which are
PT differentially expressed in different cell types

XX
PS Claim 1; Page 1729-1730; 2479pp; English.
XX
CC The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX
SQ Sequence 745 BP; 183 A; 163 C; 184 G; 192 T; 23 other;

Query Match 14.6%; Score 524; DB 20; Length 745;
Best Local Similarity 95.5%; Pred. No. 8.9e-125;
Matches 575; Conservative 0; Mismatches 22; Indels 5; Gaps 4;

QY 1 ctttttggtcactcgtggccaccggggccaaagctaggctcgatgagaagaggttg 60
Db 146 ctttttggtcactcgtggggccaccggggccaaagctaggctcgatgagaagaggttg 205
QY 61 atcctgctgtcttggaagtcgtggatctggccaaagaagtggtgacagtgacgaa 120
Db 206 atcctgctgtcttggaagtcgnggatctggccaaagaagtggtgacagtgacgaa 265
QY 121 gtgctagttagaccgagtcagttggaactgacgagagactgcaaaagaactaaata 180
Db 266 gtgctagttagaccgagtcagttggaactgacgagagactgcaaaagaactaaata 325
QY 181 gacgtcgaaagcctgctcgcgctcgagctggacccagcctccgacagtttaaccag 240
Db 326 gacgtcgaaagcctgctcgcgctcgagctggacccagcctccgacagtttaaccag 385
QY 241 tcagtgagcaatgaactgaattgtagtaggagacttctctgtctgtactgatggg 300
Db 386 tcagtgagcaatgaactgaattgtagtaggagacttctctgtctgtactgatggg 445
QY 301 cagcttcattgtagcagcaatctgcatctgaggtctccaaagaagaatgtactattac 360
Db 446 cagcttcattgtagcagcaatctgcatctgaggtctccaaagaagaatgtactattac 505
QY 361 gaatgcttctattctcttttgaattcttcgaaagaattcaagaatgtgacctgggtca 420
Db 506 gaatgcttctattctcttttga-cttcgaaagaattcaagaatgtgacctgggtca 564
QY 421 cctgatattgacaaactgg--acgttgccacaatgacagagatttataattttgagaaga 478
Db 565 cctgatattgacaaactggacgtttgcccacaatgacagagatttataattttgagaaga 624
QY 479 tagttcagctctcgatgagcctctcaagtt-gaagatatggggaataataatttta 537
Db 625 tagttcagctctcgatgagcctctcaagttcagttggaagatatggggaataataatttta 684
QY 538 gcaatgatttcagagccttataatcacaggttttcagatccagagagagtgattacaag 597
Db 685 gcaatgatttcagagccttataatc--canggtttcagatccngagagagtgattacaag 743
QY 598 tt 599

Db	744	tt 745	
AAH12166/C			
Id	AAH12166	standard; cDNA; 565 BP.	
XX	AAH12166;		
AC	AAH12166;		
XX	26-JUN-2001	(first entry)	
DT	Human cDNA clone (3'-primer)	SEQ ID NO:9001.	
XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	EP1074617-A2.		
PN	07-FEB-2001.		
XX	28-JUL-2000; 2000EP-0116126.		
XX	29-JUL-1999; 99JP-0248036.		
PR	27-AUG-1999; 99JP-0300253.		
PR	11-JAN-2000; 2000JP-0118776.		
PR	02-MAY-2000; 2000JP-0183767.		
PR	09-JUN-2000; 2000JP-0241899.		
XX	(HELI-) HELIX RES INST.		
PA	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
PI	WPI; 2001-318749/34.		
DR	Primer sets for synthesizing polynucleotides, particularly the 5602		
XX	full-length cDNAs defined in the specification, and for the detection		
PT	and/or diagnosis of the abnormality of the proteins encoded by the		
PT	full-length cDNAs -		
XX	Claim 3; SEQ ID 9001; 2537pp + CD ROM; English.		
PS	The present invention describes primer sets for synthesizing 5602		
XX	full-length cDNAs defined in the specification. Where a primer set		
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary		
CC	to the complementary strand of a polynucleotide which comprises one of		
CC	the 5602 nucleotide sequences defined in the specification, where the		
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination		
CC	of an oligonucleotide comprising a sequence complementary to the		
CC	complementary strand of a polynucleotide which comprises a 5'-end		
CC	sequence and an oligonucleotide comprising a sequence complementary to a		
CC	polynucleotide which comprises a 3'-end sequence, where the		
CC	oligonucleotide comprises at least 15 nucleotides and the combination of		
CC	the 5'-end sequence/3'-end sequence is selected from those defined in		
CC	the specification. The primer sets can be used in antisense therapy and		
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,		
CC	particularly full-length cDNAs. The primers are also useful for the		
CC	detection and/or diagnosis of the abnormality of the proteins encoded by		
CC	the full-length cDNAs. The primers allow obtaining of the full-length		
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and		
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to		
CC	AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632		
CC	represent oligonucleotides, all of which are used in the exemplification		
CC	of the present invention.		
XX	Sequence 565 BP; 192 A; 97 C; 102 G; 172 T; 2 other;		
XX	Query Match 14.1%; Score 505.4; DB 22; Length 565;		
XX	Best Local Similarity 95.8%; Pred. No. 4.8e-120;		
XX	Matches 541; Conservative 0; Mismatches 21; Indels 3; Gaps 2;		

Qy	2810	ttatattgtacttttttcttattgattggttttgacctttaataagagagaaattccatagtt	2869
Db	563	TTTAAATGGACTTTTTCATAAT--GATGGTGGACTTTTAAATAGAGAATTCCTAGATT	506
Qy	2870	ttaataatccagagaagtgagacaatttgacacagtgattctctagaaaaaacacactaac-	2928
Db	505	TTTAAATGCCAGAGTGAGCCATTTGAACAGGTGATTTCTAGAACCAATACACTAACT	446
Qy	2929	tgaacagaagtgaatgcttatatatattatgatagccttaaaccttttctcctaagtcc	2988
Db	445	TGACCAGAGTGAATGCTTATATATATATATATATATATATATATATATATATATAT	386
Qy	2989	ttaactgtcaaatataataaccttttaagcattagcagtaggacttagtcagcagtagactg	3048
Db	385	TTAACTGTCAAAATAATTATAACCTTTTAAAGCATAGGACTATAGTCAGCATGCTAGACTG	326
Qy	3049	agaggtaaacactgatgcaatttagaacaaggtaactgactgactgactgactgactgactg	3108
Db	325	AGAGGTAACACTGATGCAATTTAGAACAGGTACTGATGCTGCTGCTGCTGCTGCTGCTG	266
Qy	3109	ttagctgtgttattgctataaaagtgaatttagacacttagctagtagtactgctgctcat	3168
Db	265	TTAGCTGTGTTTATGCTATATAAAGTGCATATATAGACACTAGTACTGCTGCTCAT	206
Qy	3169	gtaactccaaagaaaacagagatttcatttaagtgcattgaatgttgatatttctctaaagt	3228
Db	205	GTAACCTCCAAAGAAAACAGGATTTTCAATTAAGTGCATTTGAATGTGGCTATTTCTCTA	146
Qy	3229	actcatattgctcttctgctgaatgcaatgccgtgcagatttatgagcgtctattttta	3288
Db	145	ACTCATATTGCTTTTGGCTTGAATGCAATGCCGTGCAATTTATGTGCTGCTATTTTAA	86
Qy	3289	ttctctgctacttcttaacaccttaagggagagaagcaaacatttctctcagctga	3348
Db	85	TTTTCTGTGCATTTACTTTAACACCTTAAAGGGAGAGAACATTCCTTTCTTCAGCTGA	26
Qy	3349	ctggcaatggccctttaactgcaat	3373
Db	25	CTGGCAATGGCCCTTTAACTGCAAT	1
RESULT	11		
AAF68133			
ID	AAF68133	standard; cDNA; 503 BP.	
XX	AAF68133;		
AC	AAF68133;		
XX	12-APR-2001	(first entry)	
DT	Human lung tumour protein related nucleotide sequence	SEQ ID NO:51.	
XX	Human; lung cancer; lung tumour; lung tumour protein; gene therapy;		
KW	lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;		
KW	cytostatic; antisense inhibition; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO200100828-A2.		
PN	04-JAN-2001.		
PD	30-JUN-2000; 2000WO-US18061.		
XX	30-JUN-1999; 99US-0346492.		
PR	15-OCT-1999; 99US-0419356.		
PR	17-DEC-1999; 99US-0466867.		
PR	30-DEC-1999; 99US-0476300.		
PR	06-MAR-2000; 2000US-0519642.		
PR	22-MAR-2000; 2000US-0533077.		
PR	10-APR-2000; 2000US-0546259.		
PR	27-APR-2000; 2000US-0560406.		
PR	05-JUN-2000; 2000US-0589184.		

XX PA (CORI-) CORIXA CORP.
XX PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
XX PI Retter MW, Mannion J;
XX XX
XX WPI; 2001-071488/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
XX useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 168; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
XX associated proteins (I) and the nucleic acids (NAs) that encode them.
XX (I) have cytostatic activity and can be used in gene therapy, antisense
XX inhibition and in vaccines. The NAs and the lung tumour-associated
XX proteins they encode may be used in the prevention, treatment and
XX diagnosis of diseases associated with their inappropriate expression,
XX especially lung cancers. For example, the NAs may be administered to
XX treat diseases by rectifying mutations or deletions in a patient's genome
XX that affect the activity of the protein by expressing inactive proteins
XX or to supplement the patients own production of (I). Additionally, the
XX NAs may be used to produce the lung-tumour associated protein, according
XX to standard recombinant DNA methodology. Conversely, antisense NA
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own genes and preventing their expression. The NA
XX and complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar NA sequences in
XX samples, and hence which patients may be in need of treatment for lung
XX cancer. The (I) may be used as antigens in the production of antibodies
XX and in assays to identify modulators (agonists and antagonists) of the
XX expression and activity of the protein. AAF68083 to AAF68878 and
XX AAB76848 to AAB76878 represent human lung tumour protein related
XX nucleotide and protein sequences which are used in the exemplification
XX of the present invention.
XX
XX Sequence 503 BP; 188 A; 69 C; 78 G; 168 T; 0 other;

Query Match 13.9%; Score 498.4; DB 22; Length 503;
Best Local Similarity 99.8%; Pred. No. 2.9e-118;
Matches 499; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2572 gatacttattgataaaacaaataattttaaaacacctgaagataaaattagaagaaa 2631
Dy 1 gatacttattgataaaacaaataattttaaaacacctgaagataaaattagaagaaa 60
Qy 2632 ttgtgacccctccacaaacatacaaaagttaaagtttgatctttttcagcaggtta 2691
Dy 61 ttgtgacccctccacaaacatacaaaagttaaagtttgatctttttcagcaggtta 120
Qy 2692 tcagttgataataatgaattaggcccaaaatgcaaaacgaaataaagaagcagctacatg 2751
Dy 121 tcagttgataataatgaattaggcccaaaatgcaaaacgaaataaagaagcagctacatg 180
Qy 2752 tagttagtaattcttagttgaactgaattgaattgcttgcctcatgtattatttt 2811
Dy 181 tagttagtaattcttagttgaactgaattgaattgcttgcctcatgtattatttt 240
Qy 2812 atattgtactttttcattattgtattgtgtgactttaataagaagaaattccatagtttt 2871
Dy 241 atattgtactttttcattattgtattgtgtgactttaataagaagaaattccatagtttt 300
Qy 2872 taatatccagaagtgaacatttgaaacagtgattttatagataaaacatacaactaactga 2931
Dy 301 taatatccagaagtgaacatttgaaacagtgattttatagataaaacatacaactaactga 360
Qy 2932 acagaagtgaatgcttatatatattatgatagccttaaacctttttccttaatgcctta 2991
Dy 361 acagaagtgaatgcttatatatattatgatagccttaaacctttttccttaatgcctta 420
Qy 2992 actgtcaataattataacaccttttaaaagcataggactatagtcagctagctagactgaga 3051

Db 421 actgtcaataataataacaccttttaaaagcataggactatagtcagctagactgaga 480
Qy 3052 ggttaaacactgatgaattta 3071
Dy 481 ggttaaacactgatgaattta 500
RESULT 12
AAF68134
ID AAF68134 standard; cDNA; 503 BP.
XX
XX AAF68134;
XX
XX 12-APR-2001 (first entry)
XX Human lung tumour protein related nucleotide sequence SEQ ID NO:52.
XX Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
XX lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
XX cytostatic; antisense inhibition; ss.
XX Homo sapiens.
XX WO200100828-A2.
XX
XX 04-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18061.
XX
XX 30-JUN-1999; 99US-0346492.
XX 15-OCT-1999; 99US-0419356.
XX 17-DEC-1999; 99US-0466867.
XX 30-DEC-1999; 99US-0476300.
XX 06-MAR-2000; 2000US-0519642.
XX 22-MAR-2000; 2000US-0533077.
XX 10-APR-2000; 2000US-0546259.
XX 27-APR-2000; 2000US-0560406.
XX 05-JUN-2000; 2000US-0589184.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
XX Retter MW, Mannion J;
XX
XX WPI; 2001-071488/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
XX useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 168; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
XX associated proteins (I) and the nucleic acids (NAs) that encode them.
XX (I) have cytostatic activity and can be used in gene therapy, antisense
XX inhibition and in vaccines. The NAs and the lung tumour-associated
XX proteins they encode may be used in the prevention, treatment and
XX diagnosis of diseases associated with their inappropriate expression,
XX especially lung cancers. For example, the NAs may be administered to
XX treat diseases by rectifying mutations or deletions in a patient's genome
XX that affect the activity of the protein by expressing inactive proteins
XX or to supplement the patients own production of (I). Additionally, the
XX NAs may be used to produce the lung-tumour associated protein, according
XX to standard recombinant DNA methodology. Conversely, antisense NA
XX molecules may be administered to down regulate protein expression by
XX binding with the cells own genes and preventing their expression. The NA
XX and complementary sequences may also be used as DNA probes in diagnostic
XX assays to detect and quantitate the presence of similar NA sequences in
XX samples, and hence which patients may be in need of treatment for lung
XX cancer. The (I) may be used as antigens in the production of antibodies
XX and in assays to identify modulators (agonists and antagonists) of the
XX expression and activity of the protein. AAF68083 to AAF68878 and
XX AAB76848 to AAB76878 represent human lung tumour protein related

